Determinants of New Technology-Based Firms Performance in Catch-Up Regions: Evidence from the U.S. Biopharmaceutical and IT Service Industries 1996-2005

A Dissertation
Presented to
The Academic Faculty

By

Wenbin Xiao

In Partial Fulfillment
Of the Requirements for the Degree
Doctor of Philosophy in Public Policy

Georgia Institute of Technology
And
Georgia State University
April, 2008

Determinants of New Technology-Based Firms Performance in Catch-Up Regions: Evidence from the U.S. Biopharmaceutical and IT Service Industries 1996-2005

Approved by:

Dr. Philip Shapira, Advisor School of Public Policy *Georgia Institute of Technology*	Dr. Greg Lewis School of Policy Studies *Georgia State University*
Dr. Marco Ceccagnoli School of Management *Georgia Institute of Technology*	Dr. John Walsh School of Public Policy *Georgia Institute of Technology*
Dr. Mary Frank Fox School of Public Policy *Georgia Institute of Technology*	Date Approved: December 4, 2007

[To my wife Jinghui Bao and daughter Ressie Xiao]

TABLE OF CONTENTS

LIST OF TABLES VI
LIST OF FIGURES VII
LIST OF SYMBOLS AND ABBREVIATIONS VIII
SUMMARY IX
CHAPTER 1: INTRODUCTION AND OVERVIEW 1
1.1. The Role of NTBFs in Regional Transformation 1
1.2. Industrial Cluster Size, Catch-up Regions, and NTBF Performance 3
1.3. Overview of the Dissertation 6
CHAPTER 2: THEORY AND HYPOTHESES 9
2.1. Location Decisions of Technological Entrepreneurs 9
2.2. Sources of Radical Innovations 13
2.3. Conceptual Framework 17
CHPATER 3: METHODOLOGY 19
3.1. Research Setting 19
3.2. Measures and Data 27
3.3. Models 42
CHAPTER 4: BIOPHARMACEUTICAL INDUSTRY 51
4.1. The Biopharmaceutical Industry 51
4.2. IPO and M&A Activities in the Biopharmaceutical Industry 58
4.3. Determinants of Regional Variations in Biopharmaceutical NTBF performance 61
4.4. Summary 83
CHAPTER 5: IT SERVICE INDUSTRY 84
5.1. The IT Service Industry 84
5.2. IPO and M&A Activities in the IT Service Industry 89
5.3. Determinants of Regional Variations in the IT Service NTBF Performance 93
5.4. Summary 111
CHAPTER 6: CONCLUSIONS 112
6.1. Summary of Main Findings 112
6.2. Policy Implications 115
6.3. Limitations and Future Research Directions 118

APPENDIX 1: LIST OF LEADING AND TOP 20 CATCH-UP REGIONS 120
APPENDIX 2: INDUSTRIAL SHARE BY STATE, IN 1995, 2000, AND 2004 122
APPENDIX 3: RESULTS OF REDUCED ZINB MODEL SPECIFICATIONS 126
REFERENCES 120

LIST OF TABLES

Table 1: Statistics of Academic R&D Expenditures by Field, 1995 and 2000 32
Table 2: Definition of Variables 36
Table 3: Occupational Job Titles for Life Scientist and Computer Scientist 38
Table 4: Flags, Range, and Midpoints in the County Business Pattern Data 40
Table 5: Supplier- and Buyer- Industries of the Biopharmaceutical Industry 42
Table 6: Supplier- and Buyer- Industries of the IT Service Industry 42
Table 7: Descriptive Statistics of the Dependent Variable 46
Table 8: Number of Zero and Non-Zero Observations for the Dependent Variable 46
Table 9: U.S. Biopharmaceutical Industry Statistics: 1995-2005 53
Table 10: Share of Total Establishments by Firm Size Classes for Biopharmaceutical Industry and National Average: 1995-2005 54
Table 11: Twenty Largest Biopharmaceutical Metropolitan Areas in 1995 and 2004 57
Table 12: Share of Biopharmaceutical Industry for Leading and Catch-Up Regions, ... 57
Table 13: Statistics of Biopharmaceutical IPO and M&A Events 60
Table 14: Descriptive Statistics for Biopharmaceutical Cross-Sectional Variables 65
Table 15: Descriptive Statistics for Biopharmaceutical Panel Data Variables 65
Table 16: Correlation Coefficient Matrix for the Biopharmaceutical Cross-Sectional Data 66
Table 17: Correlation Coefficient Matrix for the Biopharmaceutical Panel Data 67
Table 18: Results of Cross-Sectional ZINB Models of Biopharmaceutical IPO and M&A Events between 1996 and 2005 73
Table 19: Results of Two-Period Panel ZINB Models of Biopharmaceutical IPO and M&A Events between 1996 and 2005 77
Table 20: Distance-Weighted Results of ZINB Models of Biopharmaceutical 81
Table 21: U.S. IT Service Industry Statistics: 1995-2005 85
Table 22: Share of Total Establishments by Firm Size Classes for IT Service Industry and National Average: 1995-2005 86
Table 23: Twenty Largest IT Service Metropolitan Areas in 1995 and 2004 88
Table 24: Share of IT Service Industry for Leading and Catch-Up Regions, ... 89
Table 25: Statistics of IT Service IPO and M&A Events 92
Table 26: Descriptive Statistics of IT Service Cross-Sectional Variables 96
Table 27: Descriptive statistics for variables in the IT service panel data models 96
Table 28: Correlation Coefficient Matrix for the IT Service Cross-Sectional Data 97
Table 29: Correlation Coefficient Matrix for the IT Service Panel Data 98
Table 30: Results of Cross-Sectional ZINB Models of IT Service IPO and M&A Events between 1996 and 2005 103
Table 31: Results of Two-Period Panel ZINB Models of IT Service IPO and M&A Events between 1996 and 2005 106
Table 32: Distance-Weighted Results of ZINB Models of IT Service 109

LIST OF FIGURES

Figure 1: Conceptual Framework 18
Figure 2: Leading and Catch-Up Regions in the Biopharmaceutical Industry, 1995 25
Figure 3: Leading and Catch-Up Regions in the IT Service Industry, 1995 26
Figure 4: Number of Total and Sample Biopharmaceutical IPOs, 1996- 2005 59
Figure 5: Number of Total and Sample Biopharmaceutical M&A, 1996- 2005 59
Figure 6: Spatial Distribution of Biopharmaceutical IPO and M&A Events in Catch-Up Regions 61
Figure 7: Number of Total and Sample IT Service IPO Events, 1996- 2005 90
Figure 8: Number of Total and Sample IT Service M&A Events, 1996- 2005 91
Figure 9: Spatial Distribution of IT Service IPO and M&A Events in Catch-Up Regions 92

LIST OF SYMBOLS AND ABBREVIATIONS

NTBF	New Technology based Firms
IPO	Initial Public Offering
M&A	Merger and Acquisition
NBRM	Negative Binomial Regression Model
ZINB	Zero-Inflated Negative Binomial Regression Model
R&D	Research and Development
OLS	Ordinary Least Square
IT	Information Technology
VC	Venture Capital

SUMMARY

This study investigates the impacts of regional characteristics on the early-stage performance of New Technology-Based Firms (NTBFs) in catch-up regions where a mature industrial cluster has yet to be formed. It hypothesized that the average NTBF performance in a region is a function of its scientist job market conditions, cultural diversity, venture capital, academic research, industrial structure, and local entrepreneurial climate. Using the events of Initial Public Offerings (IPO) and Merger & Acquisitions (M&A) as an indicator of early-stage success of NTBFs, this study constructs a set of Zero-Inflated-Negative-Binomial (ZINB) models to predict the spatial distribution of such events in the U.S. biopharmaceutical and Information Technology (IT) service industries during the period from 1996 to 2005.

Several empirical findings emerge from this study. First, the local entrepreneurial climate plays a significant and positive role on NTBF performance in both industries. Second, the positive impact of cultural diversity is more significant in the IT service industry than in the biopharmaceutical industry. Third, the scientist job market size and absolute salary level have positive impacts on NTBF performance, but the effect of relative salary level is negative. Fourth, proximity to venture capital firms has positive but non-linear effects, but the adverse effect of excess venture capital is stronger in the IT service industry. Fifth, there is little evidence of the direct effects of academic research in determining the NTBF performance in both industries. Finally, industrial specialization is significant and positive only in the IT service industry. The results suggest that promoting local entrepreneurial climate and cultural diversity are two effective policy instruments for catch-up regions to foster their NTBF growth.

CHAPTER 1: INTRODUCTION AND OVERVIEW

The structure of the local environment not only influences individual entrepreneurs' desire to undertake risky ventures, but also affects their ability to perceive opportunities and mobilize resources (Thornton and Flynn 2003). The presence of industrial clusters, or the geographical concentration of interrelated industries, is one of the most striking features of high-tech industries (Porter 1990). In the U.S. context, the bulk of high tech firms cluster in the Silicon Valley, Boston 128, Seattle, Austin, and San Diego, and the rest distribute sparsely across the country. This dissertation explores the conditions and mechanisms that enable new technology-based firms (NTBFs) to emerge and thrive in catch- up regions, where a mature industrial cluster has yet to be formed. In particular, this study investigates what location-specific characteristics can enhance a catch-up region's attractiveness to entrepreneurial talents and its capacity of conducting radical innovations which could lead to brand new products, service, or industry. This chapter consists of three sections. Section 1.1 discusses the role of NTBFs in regional transformation. Section 1.2 summarizes a variety of views regarding the relationship between the size of industrial cluster and firm performance, and how this study contributes to the extant literature. The last section 1.3 presents an overview of this study.

1.1. The Role of NTBFs in Regional Transformation

In this study, NTBFs refer to a subset of small firms that are newly established, independently operated, and focused on commercializing technologies developed in public or private labs (Bollinger et al, 1983; Storey &Tether, 1998). These firms are distinguished from the majority of small businesses by aggressively pursuing new and innovative products or services (Acs and Storey 2004; NGA 2004). In the face of intensifying competition from newly emerging low-cost countries and widespread outsourcing practice of large companies, recent academic literature has placed increasing emphasis on fostering "homegrown" NTBFs as a sustainable strategy to reinvigorate a regional economy (Feldman and Francis 2004). In the meantime, state and local governments launched various policy initiatives to foster "homegrown" NTBFs in one or more high-tech fields with a belief that each location has a unique industrial heritage and therefore can create a high tech cluster of niche capabilities. For example, from 2001 to 2004, the number of states that were targeting the bioscience for development increased from 14 to 40, and now all 50 states have technology-based economic development initiatives that are available to bioscience companies (Biotechnology Industry Organization, 2005). Other newly targeted technology fields include but are not limited to nanotechnology, stem cell, and fuel cell. The origins of this considerable interest in NTBFs can, in some part, be traced to these firms' unique nature of innovation and their role in sparking regional transformation.

NTBFs are those founded by independent entrepreneurs and inventors, and thus provide a unique organizational environment conducive to radical innovation. Schumpeter (1934, p. 66) observed that innovations are embodied in "new firms which generally do not arise out of the old ones but start producing beside them." Utterback and Abernathy (1979) argue that the character and atmosphere of NTBFs is particularly suited to encouraging major product innovations. To compete against the existing large firms, NTBFs basically have to respond to market needs with high performance products or services, and thus place a high priority on product innovation as a competitive strategy. A study commissioned by the US Small Business Administration (1995) shows that a critical share of innovative breakthroughs of the past two centuries came from small and newly founded ventures and their affiliated entrepreneurs. More recently, Baumol (2005) argues that the market mechanism assigns the search for radical inventions to the NTBFs and their subsequent development to the large firms. Therefore, this division of innovative efforts between small and large firms is not transitory. He further points out that the existence of this type of specialization does not suggest that large enterprises are inefficient or ineffective innovators. Rather, it is the combined work of the two together that makes possible the ultimate success of technological commercialization.

From the perspective of regional economic development, a relevant policy question is whether and how NTBFs can emerge and thrive in a catch-up region characterized by "low levels of clustering, a weak endowment of relevant institutions, a lack of interaction and of networks" (Todtling & Trip, 2005, p. 1204). A region's economy often changes in an incremental manner that is heavily shaped by the established institutional and industrial structure (Fuchs & Shapira, 2005). If NTBFs can achieve superior performance even in isolated locations lacking a critical mass of industrial cluster, these firms can function as key drivers in sparking regional transformation or even breakthrough. Such events can help those catch-up regions diverge from the traditional declining industries and enter into more lucrative high technology industries or other development opportunities (Fuchs & Shapira, 2005). Otherwise, these lucrative industries will continue to be geographically concentrated in only a small number of locations, and the regional disparities in economic performance will tend to grow.

1.2. Industrial Cluster Size, Catch-up Regions, and NTBF Performance

The definition of "catch-up regions" used in this study is industry-specific. A region might take a leading position in one industry but not in others. The recent literature involves an ongoing debate on the possibility of NTBFs to achieve superior performance in a catch-up region. The conventional wisdom suggests that firms located in or near clusters of similar firms perform better than their counterparts in sparse areas (Marshall 1920; Porter 1990; Krugman 1991). Marshall (1920) identified three types of positive externalities stemming from agglomeration. First, the concentration of several firms in a single location offers a pooled market for highly skilled human capital, ensuring a lower probability of human capital shortage. Second, firm clustering creates a sufficient demand that enables local suppliers to provide highly specialized products and services, and lower the transportation costs. Third, clustered firms can benefit from localized knowledge spillovers. Following Marshall's trinity of reasons for co-location, Porter (1990; 1998) posits that firms especially benefit from the clustering of interrelated industries and supportive institutions. Krugman (1990) argues that the concentration of manufacturing in a limited number of sites mainly results from pecuniary externalities like minimizing transportation costs associated with either demand or supply linkages. In addition, the entrepreneurial culture that advocates norm-breaking behavior is particularly strong in mature industrial clusters and thus leads to relatively high entrepreneurial activities in these regions (Staber, 2005, in Fuchs & Shapira, 2005). An implicit assumption under these arguments is that the magnitude of positive externalities increases with the size of clustering.

On the other hand, several recent empirical studies question the conventional wisdom that proximity to a leading mature industrial cluster enhances average firm performance. Sorenson and Audia (2000) examine the founding and failure rates of shoe manufacturing plants in the US from 1940 to 1989. Their study suggests that while new firms typically emerge in the vicinity of major industrial clusters, firms located in concentrated regions of shoe manufacturing experienced substantially higher failure rates than did isolated firms. Therefore, they posit that locating in the large clusters adversely affects firms' chance to survive and thrive because the concentration of structurally equivalent firms increases the extent of competition for valuable inputs. Stuart & Sorenson (2003) conduct a similar study in the U.S. biotechnology industry, and find that although entrepreneurs most frequently began new biotech firms in route 128 of Boston, Massachusetts and Palo Atlo, California, these areas offer the worst location for new venture performance. According to their performance models, firms perform best in the tri-state area, where New Jersey, Pennsylvania, and New York come together (p. 248). More recently, Folta et al. (2006) use various performance indicators to examine the relationship between cluster size and firm performance in the US biotech industry. Their findings indicate the diseconomies of agglomeration seem to dominate the economic benefits of agglomeration when clusters exceed about 65 firms.

Other lines of research also indicate that firms in catch-up regions may succeed. Dumais et al (2002) examine the dynamics of geographic concentration in the US manufacturing industries between 1972 and 1992. Their results suggest that while the overall agglomeration level of the U.S. manufacturing industries has only declined slightly in the last twenty years, there is a substantial degree of variation in the locations of these agglomerations, which was partially caused by the successful startup activities in isolated locations. Florida (2002a, 2002b) posits that leading industrial clusters provide a necessary but insufficient condition to attract highly creative talents (human capital), and other regional factors like regional tolerance matter more. Therefore, regions without mature industrial clusters may still be attractive to creative talents.

In sum, previous studies don't entirely exclude the possibility for NTBFs in isolated locations to achieve superior performance. However, they leave the following questions unanswered:

First, why do some catch-up regions have higher average-NTBF-performance than do the leading mature industrial clusters? Are there any commonalities among these "successful" catch-up regions?

Second, over time, why do some catch-up regions successfully become new hotbeds of technological entrepreneurs while others continue to lag behind?

Third, besides the industrial cluster size, what other location-specific characteristics enhance the likelihood that a NTBF locating in a catch-up region will thrive?

Fourth, how do these causality patterns vary across the life cycle of specific industries? Does the biopharmaceutical industry offer more opportunities for NTBFs in the isolated locations to catch up or even leapfrog than the IT service industry, or vice versa?

1.3. Overview of the Dissertation

This dissertation examines the spatial distribution of the US NTBF Initial Public Offerings (IPO) and acquisition activities from 1996 to 2005. Following the literature on the locational choices of the "creative class" (Florida 2002a; 2002b; 2003; Lee, Florida et al. 2004) and origins of dynamic externalities (Glaeser, Kallal et al. 1992; Feldman and Audretsch 1999; Ketelhohn 2002), the present research hypothesizes that the average firm performance in a catch-up region is positively associated with the region's attractiveness to technological entrepreneurs and its capacity of conducting radical innovations, which, in turn, are determined by the local job market conditions for scientists, the availability of venture capital, cultural diversity, industry related academic research, industrial structure, and entrepreneurship dynamics.

This dissertation makes several contributions to the technology-based economic development literature. First, in contrast to many of the previous studies that emphasize the "best practice" in the leading mature clusters, this study focuses on catch-up regions. The main interest of this research is to investigate why the average NTBF performance of some catch-up regions is better than others and how the spatial heterogeneity of NTBF performance varies over time and industry. As some scholars point out, looking at a fully mature industrial cluster reveals little prescriptive information about how such regions actually develop (Feldman, Francis et al. 2005; Tödtling and Trippl 2005).

Second, this study develops a comprehensive theoretical model to explain the spatial heterogeneity in NTBF performance. Prior research has demonstrated a disproportionate focus on the impact of the industrial cluster size on the NTBF average performance in a region, but has been short of explanations on why the leading mature clusters in some industries are not associated with the highest average-NTBF-performance. This study investigates such a paradox by positing that a region's NTBF performance is a function of its attractiveness to technological entrepreneurs and its capacity of conducting radical innovations, which in turn, are determined by the local labor market conditions for scientists and engineers, venture capital endowment, cultural diversity, academic research, industrial structures, and entrepreneurship dynamics. While the first three factors mainly affect the location decisions of technological entrepreneurs, the rest three determine the capacity of conducting radical innovations. The size of industrial clusters in a region is only positively associated with some dimensions of these factors.

Third, this study examines the most recent Initial Public Offerings and acquisition activities in the US over the past decade. Unlike many of the existing studies that approximated firm performance with survival, this study views the event of IPO and acquisition as the indicator of firm early-stage success. Folta et al. (2006) noted that survival is a biased measure of firm performance because entrepreneurs in larger clusters have greater access to alternative business opportunities and therefore are more likely to terminate business for a given level of performance. Both IPO and acquisition are a successful outcome for a NTBF started to develop an innovative idea, so the spatial distribution of these events reveals where the most successful NTBFs are located.

Finally, this analysis accounts for detailed industry-specific conditions. This research focus on the biopharmaceutical and IT service industries, because they are widely-accepted high tech industries, and are most active in terms of IPO and acquisition activities over the past decade. In addition, these two industries vary in the degree of geographical concentration, the closeness to academic research, and the stage of industrial life cycle, and barriers for entry, which provide rich opportunities to investigate the impacts of industry-specific characteristics on the relationship between location and firm performance.

CHAPTER 2: THEORY AND HYPOTHESES

This dissertation is built upon two streams of literature: location decisions of entrepreneurial talents and the sources of radical innovation. It assumed that location affects firm performance in high-tech industries mainly by clustering technological entrepreneurs and facilitating the conduction of radical innovation. Section 2.1 discusses the competing theories aimed to explain the location behavior of technological entrepreneurs and then presents three hypotheses in this regard. Section 2.2 derives four hypotheses regarding the underlying mechanisms of radical innovations. A theoretical framework is presented in the section 2.3.

2.1. Location Decisions of Technological Entrepreneurs

Technological entrepreneurs play a central role in a firm's early-stage viability by identifying profitable opportunities and mobilizing needed resources from their immediate business environment. They may also utilize their network ties to outside prominent organizations to overcome inherent regional disadvantages, and access remote financial capital, skilled employees, and non-codified knowledge (Stuart and Sorenson 2005). Therefore, the uneven distribution of entrepreneurial talent greatly contributes to the regional variation in the average early-stage performance of NTBFs.

Studies of the individual characteristics of technological entrepreneurs demonstrate that level of educational attainment amongst founders of NTBFs is significantly higher than that of the working population as a whole, or of founders of other types of new businesses (Storey and Tether 1998). NTBFs that are exploiting advanced technologies require entrepreneurs who understand the leading edge technologies. This is particularly true for the biotechnology industry, in which NTBF founders typically have a science- or technology- based PhD. Other studies found that creative entrepreneurs also tend to be associated with idiosyncratic personalities and unorthodox ideas (Florida 2002a). In addition, some literature suggests that new immigrants are more likely to be entrepreneurs because they bring new ideas and cultures, are risk takers, and enjoy better access to capital through family or ethnic networks than others (Saxenian 1999; Lee, Florida et al. 2004). For example, Saxenian (1999) found that extensive networks of Chinese and Indian immigrants help people start new firms by providing contacts and financial support in Silicon Valley. In sum, technological entrepreneurs are not those who take self-employment as the last resort, but are typically highly educated individuals who possess resources, are economically mobile, and can exercise considerable choice in their locations (Florida 2002a). Places attract potential technological entrepreneurs through three interrelated mechanisms: labor market conditions, cultural diversity, and local attachments.

2.1.1. Scientist job market conditions

The traditional view offered by economists is that place attracts potential entrepreneurs by offering secure and high-paid job opportunities. Early in the 1920s, Marshall observed the "labor market pool effect" stating that individual workers could minimize their economic risk by locating in a place with many possible employers of their specialized skills. A concentration of similar firms would attract, develop, and benefit from a pool of labor with a common set of skills. Technology workers in a volatile labor market do not want to remain in a small labor market in which alternative job opportunities are few. More recently, Porter advocates this "labor market pooling" effect in his "Industrial Cluster" theory (Porter 1998; 2000). Porter (1998, p .81) argues that “because a cluster signals opportunity and reduces the risk of relocation for employees, it can also be easier to attract talented people from other locations.”

The availability of job opportunities for highly educated scientists and engineers in a region is particularly important because the potential technological entrepreneurs are mainly from those who have sufficient understanding of the leading edge technologies. As a result, regions with sufficient job opportunities for scientists and engineers are more likely to host a big pool of potential entrepreneurs than other regions. A stream of sociology literature suggests that entrepreneurs tend to be geographically inertial even though they are economically mobile (Sorenson and Audia 2000; Stuart and Sorenson 2003; Thornton and Flynn 2003). Entrepreneurs, like other people, develop geographically localized networks of friends, acquaintances, and contacts. Relocation entails serous social costs in the form of breaking old ties and making new ones (Sorenson and Audia 2000). Therefore, entrepreneurs usually want to stay in a place where they have local attachments generated from family connections, education, and work experience.

Hypothesis 1a: The size of local scientist job market positively affects the average NTBF performance in a region.

Hypothesis 1b: The average salary level of local scientists increases the average NTBF performance in a region.

2.1.2. Venture capital

A considerable literature affirms that venture capital is important to the birth and growth of NTBFs. For example, Shane and Stuart (2002) have shown that linkages with venture capitalists were more likely to lead to success, especially as it related to reaching an IPO stage. Venture capital firms can not only provide critical risk capital to a NTBF, but also serve as business management experts. Since business mentoring requires frequent face-to-face contacts, NTBFs' physical proximity to venture capitalists can improve the efficiency and effectiveness of such activities. As a result, local venture capital might serve as a talent magnet, attracting highly-skilled potential entrepreneurs from outside. Other studies, however, argue that venture capitalists have only a moderate impact on the location decisions of technological entrepreneurs for two reasons. First, venture capitalists have strong incentives to hunt for outside high-quality NTBFs which have demonstrated significant growth potential. Second, as pointed out by Bygrave (2004), most small high-techs are never likely to be VC targets, so self-financing and informal investment are far more important than venture capitals. Bygrave (2004) claims that "if self-financing and informal investment dried up, entrepreneurship would wither and die. On the other hand, if classic venture capital dried up, entrepreneurship in general would continue to flourish." [1] Finally, the excess amount of venture capital in a region is likely to flow straight to low-quality entrepreneurship(Acs and Storey 2004; Venkataraman 2004).

Hypothesis 2: The number of venture capital firms located in a region increases its average-NTBF-performance at a decreasing rate.

2.1.3. Cultural diversity

Florida argues that creative people typically have many job options, and they do not slavishly follow jobs to places (Florida 2002a; 2002b; Lee, Florida et al. 2004). In his "creative class" theory, he downplayed the importance of the labor market pooling effect in the location choice of creative people whose job is to "create meaningful new forms" and highlights the role of "cultural diversity" instead (Florida 2002b). Built upon Jacob's thesis that open and diverse

[1] William D. Bygrave, presentation in 2004 Globel Entrepreneurs Monitor Conference in London, http://www.gemconsortium.org/category_list.asp?cid=165

cities attract more talented people, he posits that diverse regions, which are tolerant and open to new ideas and have diverse lifestyle amenities, have distinct advantages in attracting and retaining creative people with unorthodox ideas. According to Florida, regional diversity can be measured by three indicators: the proportion of the population that is foreign born, the concentration of same-sex male unmarried partners, and the proportion of the population that is artists. Regions with high values on these indicators tend to have lower entry barriers, making it easier for people with various backgrounds to enter and stay within it, and thus are more likely to generate new combinations of creative ideas. Florida (2002a) empirically tests the relationship between regional diversity and the distribution of human capital in the US context, and finds that an area's level of diversity, as measured by the percent of gay households, is positively and significant associated with the percentage of its population with a bachelor's degree or above. Lee, Florida, and Acs (2004) find that social diversity, measured by the percentage of the population that is foreign born and the concentration of gay male couples, facilitates the influx of a particular kind of human capital that promotes innovation and accelerates information flow, leading to the higher rate of new firm formation. In sum, according to Florida and others, cultural diversity leads to the concentration of entrepreneurial talents, and, in turn, positively affects the average performance of a region's NTBFs.

Hypothesis 3: Cultural diversity increases the average NTBF performance in a region.

2.2. Sources of Radical Innovations

Radical innovations create entrepreneurial opportunities. Entrepreneurial opportunities not only come from a firm's own Research and Development (R&D) activities, but also from knowledge investments by incumbent enterprises, universities, and other research organizations. This is particularly true for NTBFs, which lack the resources of their larger counterparts and are more dependent on exploiting technologies underutilized by the original owners. As Audretsch and Lehmann (2005) point out, without knowledge spillovers, NTBFs should have not generated high innovative output given their negligible R&D expenditures. Knowledge spillover refers to the fact that a firm is able to freely exploit economically useful knowledge produced by other agents. The extant literature offers competing theories on how the magnitude of localized knowledge spillovers in a region is affected by the local academic research, industrial structures, and entrepreneurial climate.

2.2.1. Academic research

Research universities are a major source of new knowledge and radical innovation. New knowledge generated from basic and applied research conducted at the university spills over, and fuels the innovative activities of private firms. The studies of Jaffe et al (1993), Audretsch and Feldman(1996) and others provided evidence concerning the spatial dimension of knowledge spillovers. They suggest that knowledge spillovers tend to be geographically bounded within the region where the new economic knowledge was created. Geographic proximity to the original institutes helps a firm lower the cost of accessing and absorbing knowledge spillovers. This is particularly crucial for tacit knowledge, which requires face-to-face interaction to be communicated.

Other studies, however, argue that strong network ties rather than geographic proximity to the knowledge sources affect the magnitude of knowledge spillovers (Stuart and Sorenson 2005). Bee (2003) finds that proximity to universities with leading-edge research does not appear to bestow geographic advantages to local companies, at least in the case of semiconductors. World-class universities like MIT and Stanford have worldwide networks. Companies on the other side of the world are as likely to access their cutting-edge research as local companies. His empirical analysis of blockbuster patents in semiconductors shows that technology spin-outs from major universities are not geographically concentrated in the region or state of origin. Similarly, Malmberg & Power (2005) argue that the global rather than local linkages may be the most useful method of boosting regional innovativeness, and firms can offset local factor disadvantages by building relations with key strategic partners globally.

So, assuming proximity to knowledge source is critical to the magnitude of localized knowledge spillover, we can develop the following hypothesis:

Hypothesis 4: Industry-specific academic research has a positive impact on the average NTBF performance in a region.

2.2.2. Industrial specialization

Whether industrial specialization (the degree to which a location specializes in one industry) or diversity (the range of different industries in a location) of economic activity better promotes knowledge spillovers and subsequent firm performance has been the subject of a heated debate in the economic literature (Glaeser, Kallal et al. 1992; Feldman and Audretsch 1999; Ketelhohn 2002; Porter 2003). As characterized by Glaeser and colleagues (1992), there are three camps of competing arguments on how local industrial structures affect the magnitude of knowledge spillovers: Marshall (1920)-Arrow(1962)-Romer(1986)(MAR); Jacobs (1969); and Porter (1990, 2003).

The MAR framework maintains that the most relevant knowledge spillovers occur among firms in the same industry. Specialized locations with high-level of industry concentration should experience more innovation and faster growth. In contrast, the Jacobs framework posits that the most important knowledge spillovers take place across different industries. Jacobs' theory predicts that industries will innovate more and grow faster in locations with greater diversity. Porter's position lies between those of MAR and Jacobs, arguing that the most relevant knowledge spillovers occur among a set of related industries, which include buyer and supplier industries. Porter (2003) also questions the appropriateness of using individual industries as unit of analysis because of the externalities across related industries within clusters. He (2003) posits that "a diverse array of over-lapping clusters should be associated with better performance than a diversity of clusters that are unrelated" (p. 562)

Hypothesis 5a: The degree of industrial specialization in a region promotes its NTBFs' performance.

Hypothesis 5b: Coagglomeration with buyer-industries promotes NTBFs' performance.

Hypothesis 5c: Coagglomeration with supplier-industries promotes NTBFs' performance.

2.2.3. Local entrepreneurial climate

A region's local entrepreneurial climate can affect the magnitude of localized knowledge spillovers by promoting inter-firm employee mobility. According to Saxenian (1996), the pivotal difference in the extent of inter-firm employee mobility was one of the reasons that the electronics and computer industry in Silicon Valley experienced continuing vitality and growth in the 1970s and 1980s whereas those in Route 128 corridor in Massachusetts suffered from relative stagnation and decline. In Silicon Valley, engineers and other professionals often changed firms or quit their jobs to start firms of their own, and such decisions enjoyed widespread support. In contrast, in Massachusetts, leaving one firm for another was infrequent and frowned upon as disloyal. Saxenian (1996) claims that such regional heterogeneity brought a distinctive advantage to Silicon Valley companies over those in Route 128 because the frequent inter-firm mobility enables firms to keep strengthening and updating their technologies and competencies. Hence, the dynamic local entrepreneurial activities can facilitate knowledge spillovers and therefore increase the average performance of NTBFs.

Hypothesis 6: Entrepreneurial culture positively affects the average performance of NTBFs in a region.

2.3. Conceptual Framework

Figure 1 presents a schematic that incorporates the above ideas. It depicts the causal paths through which a set of location-specific characteristics impact the average performance of NTBFs in a region.

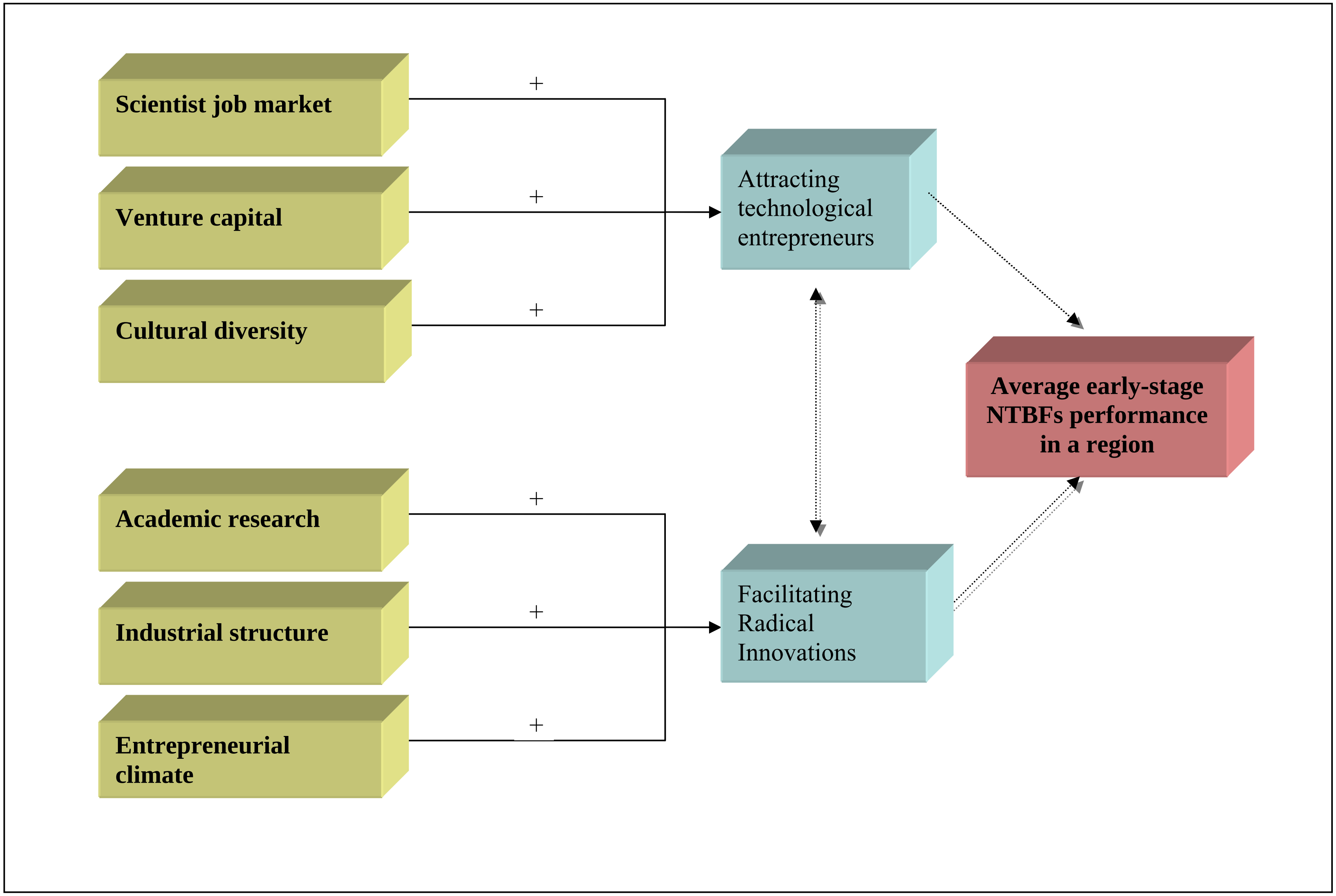


Figure 1: Conceptual Framework

CHPATER 3: METHODOLOGY

This chapter presents an empirical methodology for investigating the effects of regional characteristics on the average performance of NTBFs in the catch-up regions. Specifically, this study uses the events of Initial Public Offering (IPO) and Merger & Acquisition (M&A) as an indicator of early-stage success of NTBFs, and constructs a set of Zero-Inflated-Negative-Binomial (ZINB) models to predict the spatial distribution of such events in the biopharmaceutical and IT service industries during the period from 1996 to 2005. The first section of the chapter defines the research setting of this study. A detailed description of the variable operationalization and data sources follows in the next section. The last part of the chapter presents the model specifications.

3.1. Research Setting

3.1.1. Definition of NTBFs

NTBFs are a subset of small firms that are newly established, independently operated, and have a focus on commercializing a technology for the first time (Bollinger, Hope et al. 1983; Storey and Tether 1998). While one can usually clearly identify a firm's age and organizational independence, it is not a simple matter to measure the extent to which a firm's innovation discontinued from the extant knowledge base. According to a restricted definition, a NTBF should refer only to new independent enterprises which are developing new industries and have the potential to fundamentally transform the ways in which societies and markets operate (Storey &Tether, 1998). Other authors have embraced all new small firms operating in 'high technology' sectors (Colombo and Grilli 2005; Motohashi 2005).

Recognizing that the definition has to reflect the available data and that a narrow definition of NTBFs was not feasible, this study defines NTBFs in a broad sense. Specifically, a firm is treated as a NTBF if it meets the following criteria:

First, it has no more than 20-year-long operation history[2];

Second, it must have fewer than 500 employees, the size standard of small business defined by the US Small Business Administration;

Third, it is totally independent rather than a part or subsidiary of another incumbent firm (Bollinger et al, 1983);

Finally, it operates in a technology-based industry such as the biopharmaceutical and IT service industries.

3.1.2. Choice of industries

This study investigates the spatial heterogeneity of NTBF performance in the biopharmaceutical and IT service (computer-related) industries, subsequently. By studying a single industry subsequently instead of a cross-section of industries, it may enable us to identify industry-specific institutional factors that might explain our empirical findings (Ketelhohn 2002). The biopharmaceutical and IT service industries were selected mainly for three reasons.

First, both are among the widely-accepted high tech industries. An industry is typically considered as “high tech” if it is associated with a high proportion of research and development (R&D) expenditures and a “high” proportion of scientific, technical, and engineering personnel. However, because there are no standardized threshold values for these input-based criteria, the lists of high-tech industries produced in the literature differ from one another. These two

[2] Granstrand (1998, p. 466) argues that the newness of technology could be defined as less than 20 years, because according to the international patent system, the maximal patent lifetime is 20 years.

industries, however, are included in most high-tech industry lists (Walcott 2000; Brau, Brown et al. 2004; Hecker 2005).

Second, they differ substantially in terms of their closeness to basic science and degree of geographical concentration so that the cross-industry comparisons can enhance the validity and generalizability of the results. Prior research shows that a strong link exists between basic research mainly conducted by university scientists and the innovative performance of NTBFs in the biopharmaceutical industry (Audretsch and Stephan 1996; Deeds, Decarolis et al. 1997; McMillan, Narin et al. 2000; Lim 2004). For example, based upon their patent citation analysis, McMillan et al.(2000) found that the biopharmaceutical industry relies on basic science much more heavily than other industries and that the biopharmaceutical NTBFs rely on basic science to a much greater extent than large, diversified pharmaceutical companies do. In contrast, the innovation activities in the IT service industries are more engineering-based problem-solving that seldom reflects fundamental new knowledge. Because of the inter-industry difference in the distance to basic research, proximity to academic research should play a more important role in determining the performance of biopharmaceutical NTBFs than that of IT service NTBFs.

In terms of agglomeration, the biopharmaceutical industry is more spatially concentrated than the IT service industry. In 1996, 140 of 331 Metropolitan Statistical Areas (MSA) had no employment for the biopharmaceutical industry but all the MSAs had employment in the IT service industry[3]. The industrial agglomeration level corresponds to the degree of skewness of the spatial distribution of critical resources for the formation and growth of NTBFs (Stuart and Sorenson 2003). Thus, catch-up regions should have higher average-NTBF performance in the IT service industry than in the biopharmaceutical industry. Chapter 4 and 5 will discuss more industrial idiosyncrasies such as size, structure, barriers for new firms to enter, etc.

[3] Data source: calculated based upon the 'County Business Pattern' data of US Census Bureau.

Finally, this study measures the average performance of NTBFs in a region by counting the events of IPO and M&A between 1996 and 2005. The IPO and acquisition activities were highly active in these two industries during my study period. Thus, I can obtain a relatively large sample size by focusing upon these two industries.

This study defines the industries by a 3-digit 1987 SIC code or a 5-digit 1997 North American Industry Classification System (NAICS) code. Specifically, the biopharmaceutical industry is defined by SIC 283 prior to 1997 and NAICS 32541 thereafter. The IT Service industry corresponds to SIC 737 prior to 1997 and NAICS 5112 and 5415 thereafter. The SIC code is still in use because the current Securities Exchange Commission (SEC) Filings & Forms, which are the main data sources of public-traded companies, still categorize a firm's industrial affiliation based on the SIC code. The three-digit SIC level of analysis permits us to identify important differences that are missed at the two-digit level of aggregation. At the same time, further disaggregating to the four-digit level would sacrifice some of the coherence retained within the three-digit industry (Smith 2004). For example, in the IT service industry, there is no gain for our research purpose to make a distinction among Computer Programming Services (SIC= 7371), Prepackaged Software (SIC= 7372), and Computer Integrated Systems Design (SIC= 7373).

3.1.3 Unit of geographical observation

The units of observations in this study are the U.S. Metropolitan Statistical Area (MSA). Since the overwhelming majority of technology-based innovation and entrepreneurial activities occurred in urban areas, the metropolitan area is generally considered a more appropriate unit than a state to effectively capture the regional variation in technological entrepreneurship. For example, Feldman and Audretsh (1999) find that in the U.S., 96 percent of new innovations in 1982 originated in metropolitan areas. The U.S. Conference of Mayors and the National Association of Counties reported that 95 percent of high tech job creation between 1992 and 1999 took place in metro areas[4]. Storey and Tether (1998) observe that in Europe, most NTBFs are also located in and around the major urban areas.

One challenge to using the MSA-level data is that the Office of Management and Budget (OMB) frequently revises the definitions of MSAs to reflect the most recent Census Bureau population estimates. Consequently, the boundaries of some individual MSAs' are not stable over the past decade, which makes their historical data incomparable. Because the boundaries of sub-MSA spatial units like county or ZIP code are more stable over time than the MSA itself, this study resolved the incomparability issue by collecting most data at sub-MSA level and then aggregating them to a consistent set of MSAs.

The MSA definition used in this study was published by the OMB in March 2004 (OMB Bulletin No. 04-03) [5]. The study does not adopt more recent MSA definitions because the March 2004 definition is used by the Bureau of Labor Statistics for its present and historical Quarterly Census of Employment and Salarys (QCEW) data, which is another important data source used in this study. The March 2004 definition listed 361 MSAs in the United States. Each individual MSA has at least one urbanized area of 50,000 or more population, plus adjacent territory that has a high degree of social and economic integration with the core. In total, they cover about 83 percent of the U.S. population and 1,090 of 3141 counties.

This study excluded 49 MSAs which didn't exist in the June 1999 MSA definition. This is because data for some independent variables used in this study (i.e., scientist job market

[4] Source: http://www.usmayors.org/uscm/news/press_releases/documents/metroecon052300.htm
[5] The detailed definition is available at: http://www.whitehouse.gov/omb/bulletins/fy04/b04-03.html

conditions, proportion of foreign immigrants, and new firm birth rate) were only available upon the June 1999 MSA definition. This makes it impossible to include the 49 new MSAs into this study. The study is further limited to those MSAs where the biopharmaceutical or IT service industry consistently existed during the study period from 1995 to 2004. This guarantees that each observation in the sample has the same exposure time. In other words, we can count the IPO and M&A events that occurred in each MSA over the same period of time. In the end, 168 MSAs were identified for the biopharmaceutical industry, and 316 MSAs for the IT service industry.

3.1.4 Definition of catch-up and leading regions

This study developed a high tech index, which is a variant of the Milken Institute's 'Tech-Pole' score (DeVol 1999), to measure the relative position of an MSA in the biopharmaceutical or IT service industry. Specifically, the high-tech index is an equally weighted indicator of two factors: (1) a region's share of national total establishments in a specific high tech industry; and (2) its establishment density level, measured by the number of industrial establishments per square mile. While the first factor of the index favors large metropolitan areas, the second favors smaller regions with a relatively large number of high tech establishments. By combining them, the index creates a less biased indictor of a region's overall strength in a specific high-tech industry.

This study defines a region as a catch-up or leading region based upon its high tech index value in 1995, one year before I began to measure the regional variation in the NTBFs' early stage performance. The index is based upon the establishments instead of employment data mainly for two reasons. First, the employment and establishment variables are typically highly correlated with each other. For example, in 1995, the correlation between the number of establishment and number of employees across all the MSAs is 0.89 in the biopharmaceutical industry and 0.94 in the IT service industry. Thus, measures derived from either of them are qualitatively similar. Second, due to data disclose restrictions, many MSA regions only have rough estimates of industrial employment data. Thus, the establishment data is more accurate than the employment data for some regions.

An MSA is defined as a 'leading' region if its high tech index value in 1995 was ranked 95th percentile or above. The 95th percentile rank was chosen as the cutoff point because it can result in a fairly stable list of leading regions during the study period. Such an approach results in 8 leading regions and 160 catch-up regions in the biopharmaceutical industry. In the IT service industry, there were 15 leading regions, and 301 catch-up regions. Figure 2 and 3 depict the spatial distribution of all the leading and catch-up regions in these two industries. Appendix 1 presents the list of leading and top 20 catch up regions and their high-tech index values both in 1995 and 2000. The results indicate that in the biopharmaceutical industry, 7 of 8 leading regions were consistently ranked above 95th percentile both in 1995 and 2000. In the IT service industry, 11 of 15 leading regions were consistently ranked above 95th percentile. No leading region in 1995 was ranked below 90th percentile in 2000.

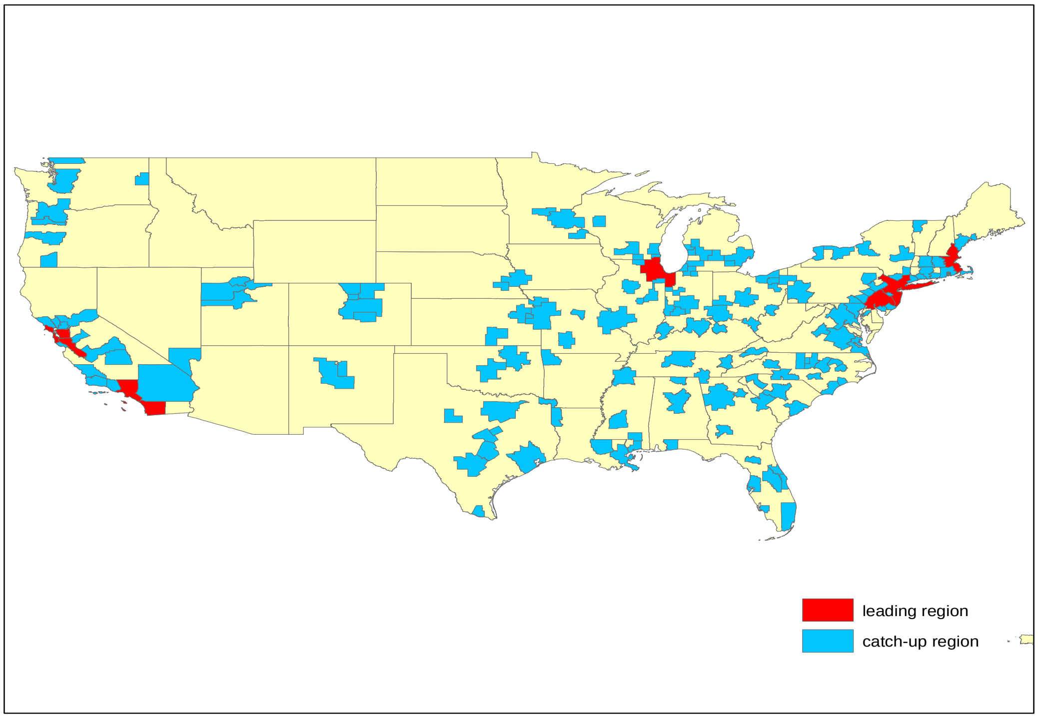


Figure 2: Leading and Catch-Up Regions in the Biopharmaceutical Industry, 1995

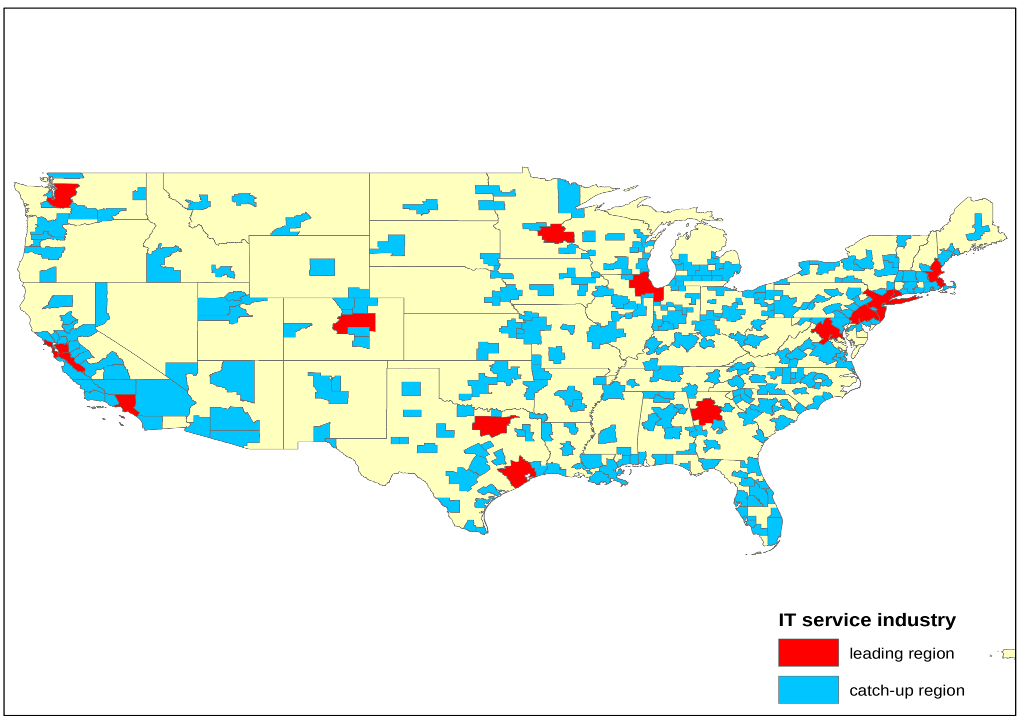


Figure 3: Leading and Catch-Up Regions in the IT Service Industry, 1995

3.2. Measures and Data

Based upon the theoretical framework developed in chapter 2, this study models the average NTBF performance in a region as a function of local scientist job market condition, proximity to venture capital, cultural diversity, academic research, industrial structure, and entrepreneurial climate. This section presents the detailed description of variable operationalization and data sources. Table 2 in the end of this section summarizes the variable definitions used in the empirical estimation.

3.2.1. Dependent variable

Prior research developed several measures of firm performance such as survival rate, asset/sales/revenues/employee growth rate, number of patents, time to IPO, etc. The survival rate over a short-period time is a performance measure with low threshold-value. Usually, we not only want to know whether firms can survive, but also how they succeed. The traditional financial indicators are of limited use for two reasons. First, NTBFs generally have few assets, revenues, or net earnings because of their limited operation history. Second, when firms are still in private status, their financial data are typically not publicly available. Patenting is one of the best available and comparable measures of a firm's innovative activity, primarily because "patent data are public, computerized, and represent a disinterested decision by an examiner that a non-obvious and potentially commercially useful invention has occurred" (Jaffe, p. 74, in Branscomb and Keller, 1998). However, since not all inventions are patentable and a firm's desire to patent varies across industry, the pitfalls of using patent-related indicators to measure a firm's success are not negligible.

This study measures the average NTBF performance in a region by counting the number of NTBF IPO and M&A events that occurred in that region during a specific period of time. This indictor of NTBF performance is based upon the fact that on average, NTBFs that have experienced an IPO or M&A event are more successful than those who have not.

The extant literature demonstrates various benefits to NTBFs by going public. First, an IPO is a critical financing tool. NTBFs usually have poor access to debt finance because of their highly variable returns, substantial information asymmetries between firms and potential investors, and limited collateral values of R&D investments (Carpenter and Petersen 2002). By going public, the issuing firm can obtain a substantial amount of capital to finance its investment opportunities. Second, going public is an organizational milestone that increases a firm's legitimacy in the business community and improves its internal management qualities. Finally, it offers an opportunity to reward the entrepreneurs and early-stage investors like venture capitalists.

Similarly, NTBF owners are very likely to be rewarded with a very high payoff through merger and acquisition because the acquirers usually have more resources and are able to offer a good deal to the targets so that they can access their technology or eliminate the potential threats. In sum, both IPOs and acquisitions can be considered a successful outcome for an NTBF started to develop an innovative idea. As a result, I can approximate the average early-stage success of a region's NTBFs by counting the number of IPO and acquisition events among them.

$\mathbf{IPO\_Target_{mt}}$ = Number of IPO and M&A events in industry i in MSA m at time t

This measure is by no means perfect, because it does not account for the fact that there is substantial variation among the IPO or acquired NTBFs in terms of their individual performance. Some IPOs fail to survive several years after they become public, and some acquired firms get much higher payoffs than others. Thus, it is necessary to compare the average performance of these "successful" NTBFs in a region to those located elsewhere. However, the data source this study uses indicates that while it is feasible to analyze the average performance of IPO firms across regions, it is not practical to do so for M&A firms. Furthermore, because IPO firms are highly concentrated in a small number of regions, only 34 MSAs had IPOs in the biopharmaceutical industry and 65 MSAs in the IT service industry during the study period. Due to such a small sample size, no independent variables turned out to be statistically significant. Therefore, I don't report the details in this study.

3.2.2. Independent variables

All the location-specific variables are defined at the MSA level. The measures of these variables are mainly drawn from prior studies (Glaeser, Kallal et al. 1992; Ketelhohn 2002; Stuart and Sorenson 2003).

(1) Scientists job market conditions

Three variables measure the local job market conditions for scientists and engineers. The first is a region's share of national industry-related scientist occupational jobs. This variable reflects the absolute size of local occupational job opportunities for industry-related scientists (i.e., life scientists for the biopharmaceutical industry and computer scientists for the IT service industry). The second variable is the average annual salary for life scientists or computer scientists. The salary rate has mixed effects on firm performance. On the one hand, the current employees, who might become entrepreneurs in the future, can be attracted by a region's higher real salary rates. As a result, these regions will be able to accumulate a pool of latent entrepreneurs. On the other hand, regions with high salary rates become less attractive for those active entrepreneurs who are creating or running a NTBF, because high real salary rates mean high business costs. The ultimate effect of salary rates on a region's NTBF performance is determined by the dominant mechanism. The third variable is calculated by dividing the average annual salary of scientist jobs by that of total occupational jobs in a MSA. This variable is constructed to adjust for the local living costs.

Job_ share$_{mt}$ = Share of national industry related scientist occupational jobs within a metro area m in time t

Absolute_salary$_{mt}$ = Ratio of annual average salary of industry related scientist occupational jobs in a metro area m in time t to the corresponding national annual average salary

Relative_salary$_{mt}$ = Ratio of annual average salary of industry related scientist occupational jobs to that of all jobs within a metro area m in time t

(2) Venture capital

Venture capitalists (VC) prefer to fund spatially proximate start-ups (Sorenson and Stuart 2001). Because the size of venture capital investment fluctuates dramatically over time, this study counts the number of VC firms located within a metro area. In addition, many VC firms are highly specialized in a specific field. Thus, it is more appropriate to count VC firms that are actively investing in the biopharmaceutical or IT service firms than just use the total number of all the VC firms. The squared term of venture capital firms is constructed to account for the adverse effects of excess amount of venture capital on firm performance as documented in some prior studies (Stuart and Sorenson 2003; Acs and Storey 2004; Venkataraman 2004).

VC_firm_{mt} = Number of industry specialized venture capital firms within a metro area m in time t

$VC\_firm\_sq_{mt}$ = Square of VC_firm_{it}

(3) Cultural diversity

Cultural diversity is a multidimensional concept and could stem from different ethnicity, religion, national origin or other characteristic(Ottaviano and Peri 2004). Florida (2002b) measures the level of cultural diversity, or the barriers to entry for diverse human capital in a region by four indices: the Gay and Lesbian index, Melting Pot Index (foreign born population), Bohemian Index (artistically creative people), and a Racial Integration Index. However, as the author acknowledged, the effects of these diversity measures vary by size of region. While the bohemian and gay measures matter much more for large regions with over 1 million people, immigration is a more appropriate indictor for small- and medium-sized regions[6].

Ottovanio & Peri (2004) define the cultural diversity index by 1 minus the Herfindal index of concentration across groups. This index reaches its maximum value 1 when each individual is in a different group and its minimum value of 0 when all individuals belong to the same group. They found a positive and high correlation coefficient between the share of foreign-born people in a city and its diversity index, ranging from 0.84 to 0.87 in the US context. This finding suggests that the presence of a large share of foreign-born, more than their group composition, is the largest source of diversity in US cities. Therefore, this study simply uses the share of foreign born in each MSA to measure its cultural diversity.

[6] Florida(2004), "Response to Edward Glaeser's review of The Rise of the Creative Class", http://creativeclass.com/rfcgdb/articles/ResponsetoGlaeser.pdf, accessed on 8/30/2007.

Immi_share$_{mt}$ = Percentage of population that were foreign born within a metro area m in time t

(4) Academic research

This study measures a region's relative strength of industry relevant academic research by its share of national academic R&D expenditure in a specific industry related field in time t.

Univ_share$_{mt}$ = Share of industry related university R&D expenditure within a metro area m in time t

The academic R&D expenditure data was compiled from the National Science Foundation's (NSF's) Survey of Research and Development Expenditures at Universities and Colleges. In each year, the NSF reports the academic R&D expenditure by fields and institutions. Life science is the most relevant field for the biopharmaceutical industry. The detailed R&D expenditure data for the top 150 institutions are available, and these institutions usually accounts for more than 96 percent of total academic R&D expenditure in the life science field (Table 1). In the computer science field, which is the most relevant field to the IT service industry, data for the top 100 institutions were reported. These institutions accounted for about 90 percent of academic R&D expenditure in the field. I identify each of top institution's ZIP code and corresponding MSA code, and aggregate the university-level expenditure data to the MSA level.

Table 1: Statistics of Academic R&D Expenditures by Field, 1995 and 2000

	Computer science		Life science	
	1995	2000	1995	2000
Total, included institutions ($)	620,045	796,293	11,679,656	16,757,635
Total, all other sampled institutions ($)	60,762	79,708	484,106	709,899
Total, all institutions ($)	680,807	876,001	12,163,762	17,467,534
Share of included institutions	91.08%	90.90%	96.02%	95.94%

Source: National Science Foundation/Division of Science Resources Statistics, Survey of Research and Development Expenditures

(5) Industry specialization

Following Glaser et al (1992) and Ketelhohn (2002), I use three traditional "location quotients" to capture the positive externalities resulted from industry specialization and co-agglomeration with the upstream and downstream industries. The location quotients are commonly used to measure the extent to which a region is more specialized in an occupation than the nation as a whole. Benchmarked at 1, a location quotient greater than 1 indicates that a MSA has a relatively high concentration of certain industries compared to the United States overall, while a value less than 1 indicates the concentration level in a MSA below the national average.

$$Ind\_lq_{mt} = \frac{\dfrac{\text{Employment in industry I in MSA m at time t}}{\text{Total employment in MSA m at time t}}}{\dfrac{\text{Total U.S. employment in industry I at time t}}{\text{Total U.S. employment at time t}}}$$

$$Buyer\_lq_{mt} = \frac{\dfrac{\text{Employment in primary buyer of industry I in MSA m at time t}}{\text{Total employment in MSA m at time t}}}{\dfrac{\text{Total U.S. employment in primary buyer of industry I at time t}}{\text{Total U.S. employment at time t}}}$$

$$Supplier\_lq_{mt} = \frac{\dfrac{\text{Employment in primary supplier of industry I in MSA m at time t}}{\text{Total employment in MSA m at time t}}}{\dfrac{\text{Total U.S. employment in primary supplier of industry I at time t}}{\text{Total U.S. employment at time t}}}$$

(6) Entrepreneurship

There are a variety of measures used in the previous studies to gauge the amount of entrepreneurial activity in an economy as well as how the entrepreneurship changes over time. For example, Becker (1984) and Blau (1987) measure the level of entrepreneurship in the U.S. between 1948 and 1980s by identifying the number of self-employed individuals using data from the Current Population Survey conducted by the Census Bureau. Reynolds (1992) and Kirchhoff and Philips (1992) focus on the rate of new firm formation using data from Dun & Bradstreet Market Identifier prepared by the Small Business Administration. Gartner and Shane (1995) argue that the number of organizations per capita is a more appropriate indicator of entrepreneurship over time then self-employment or firm formation measure.

This study focuses on the average performance of NTBF in a region, so the firm-level rather than individual-level indicator serves our purpose better. In addition, as mentioned before, due to the data disclose restriction, the establishment data is more accurate than the employment data for many MSAs. Therefore, there is no reliable data to calculate the industry-specific 'organization per capita' measure proposed by Gartner and Shane (1995). This study uses the traditional small firm birth rate to measure the overall local entrepreneurial climate.

$$\text{Firm\_Birth\_Rate}_{mt} = \frac{\text{Number of small establishments born in MSA m in year t}}{\text{Number of small establishments at the beginning of year t in MSA m}} \times 100$$

3.2.3. Control variables

The empirical estimation analysis controls for the region size, industrial cluster characteristics, and time effect that might account for the regional variation in the average performance of NTBFs.

A region's size is measured by its total employment:

$$Emp\_share_{mt} = \frac{\text{Total employment in MSA m in time t}}{\text{Total U.S. employment in time t}}$$

The industrial cluster characteristics are measured by two aspects of industrial establishments in a region: (1) number of small and medium establishments in the industry, and (2) number of industrial establishments per square mile. The establishment rather than employment data are used for the same reason discussed in section 3.1.4. That is, the employment and establishment variables are highly correlated with each other and thus measures derived from either of them are qualitatively similar. Also, the establishment data is more accurate than the employment data for some regions. The first variable reflects the absolute size of industrial cluster in a region. The second one is constructed to measure the adverse effect of 'excess' number of similar firms. The third one indicates whether firms in a region are highly concentrated or sparsely distributed.

Ind_SME_$_{mt}$ = Number of small and medium establishments in industry i within a metro area m in time t

Ind_SME sq$_{mt}$ = Ind_SME_$_{mt}$

$$Est\_density_{mt} = \frac{\text{Number of establishments in industry i in MSA m at time t}}{\text{Total land size of MSA m}}$$

The fixed time effect should be controlled in the panel data which will be discussed shortly. The capital market condition cycles from hot to cold (Ritter and Welch 2002; Helwege and Liang 2004). This is particularly true for the IPO market. The ability and intention of a firm to issue IPO is affected substantially by the varying equity market condition. Even though there is no difference in their intrinsic values, firms that issue IPOs in a "hot market" can get a higher price than those in a "cold market". This creates incentives for entrepreneurs to time their decisions to go public (Benninga, Helmantel et al. 2005). As a result, the volume of IPO within a specific period is affected substantially by the market condition. This study creates a dummy variable to control for the time effect.

Period_dummy = 0 for the period from 1996 to 2000, and 1 for the period from 2001 to 2005.

Table 2: Definition of Variables

Variable name	Definition	Time index t value	
		Model 1[a]	Model 2[b]
IPO_Target_{mt}	Number of NTBF IPO and M&A events that occurred in industry i within a metro area m in time t	1996-2005	1996-2000, 2001-2005
Job_share_{mt}	Share of national industry i related scientist occupational jobs within a metro area m in time t	1999	1999, 2004
$Absolute\_salary_{mt}$	Ratio of annual average salary of industry i related scientist occupational jobs in a metro area m in time t to the national average	1999	1999, 2004
$Relative\_salary_{mt}$	Ratio of annual average salary of industry i related scientist occupational jobs to that of all jobs within a metro area m in time t	1999	1999, 2004
VC_firm_{mt}	Number of industry i specialized venture capital firms within a metro area m in time t	1999	1999, 2004
$VC\_firm\_sq_{mt}$	Square of VC_firm_{it}	2000	2000, 2005
$Immi\_share_{mt}$	Percentage of population that were foreign born within a metro area m in time t	1990	1990, 2000
$Univ\_share_{mt}$	Share of industry i related university R&D expenditure within a metro area m in time t	1995	1995, 2000
Ind_lq_{mt}	Location quotient of industry i within a metro area m in time t	1995	1995, 2000
$Buyer\_lq_{mt}$	Location quotient of the buyers of industry i within a metro area m in time t	1995	1995, 2000
$Supplier\_lq_{mt}$	Location quotient of the suppliers of industry i within a metro area m in time t	1995	1995, 2000
$Firm\_birth_{mt}$	New small firm birth rate within a metro area m in time t	1995	1995, 2000
$Est\_density_{mt}$	Number of establishments in industry i per square mile within a metro area m in time t	1995	1995, 2000
Ind_SME_{mt}	Number of small and medium establishments in industry i within a metro area m in time t	1995	1995, 2000
$Ind\_SME\ \_sq_{mt}$	Square of Ind_SME_{it}	1995	1995, 2000
Emp_share_{mt}	Share of total national employment within a metro area m in time t	1995	1995, 2000
Period_dummy	Dummy variable indicting whether (=1) or not (=0) it is the second time period 2000-2005.		

Note: a: Cross-sectional models; b: Two-period panel data models.

3.2.4. Data

The 1996- 2005 IPO and M&A data were mainly retrieved from the Securities Data Company's (SDC) new issues database. All unit offerings, American Depository Receipts

(ADRs), and closed-end funds are excluded. I made substantial efforts to obtain a comprehensive IPO list by cross-checking the SDC data with three other sources: Hoover's IPO list (which only keep actively-traded public companies), Professor Jay Ritter's 1975-2005 IPO dataset, and University of Chicago's Center for Research in Security Prices (CRSP). Because the accuracy of IPO industry affiliation and location information is critical for this study, I obtain each company's four-digit SIC and ZIP code information from the Security Exchange Commission (SEC)'s Edgar database[7].

The occupational job data are from the Occupational Employment Statistics (OES) of the Bureau of Labor Statistics, U.S. Department of Labor. The OES survey collected both occupational employment and salary data nationwide for the first time in 1996. Since its inception, the OES program has adopted two occupational classification systems. Estimates from 1999 and subsequent years are based upon the 2000 Standard Occupational Classification (SOC) system, and therefore are not directly comparable with previous years' OES estimates. This study uses data in 1999 to measure the job market condition for life scientists and computer scientists in each MSA for the period between 1995 and 2000, and data in 2004 for the period between 2001 and 2005. Missing data are replaced with the average value between 1999 and 2004. Table 3 lists the 2000 SOC occupational job titles for scientist and engineers who develop or routinely utilize technologies that are relevant to the biopharmaceutical and IT Service industries.

[7]SEC Edgar database: http://www.sec.gov/edgar/searchedgar/companysearch.html.

Table 3: Occupational Job Titles for Life Scientist and Computer Scientist

SOC Code	SOC title
Life scientist group	
17-2031	Biomedical engineers
19-1021	Biochemists and biophysicists
19-1022	Microbiologists
19-1041	Epidemiologists
19-1042	Medical scientists except epidemiologists
19-2031	Chemists
19-4021	Biological technicians
19-4031	Chemical technicians
29-2011	Medical and clinical laboratory technologists
29-2033	Nuclear medicine technologists
29-2034	Radiologic technologists and technicians
Computer scientist group	
15-1011	Computer and information scientists
15-1021	Computer programmers
15-1031	Computer software engineers applications
15-1032	Computer software engineers systems software
15-1051	Computer systems analysts
15-1061	Database administrators
15-1081	Network systems and data communications analysts
17-2061	Computer hardware engineers
17-2071	Electrical engineers
17-3023	Electrical and electronic engineering technicians
27-1014	Multi-media artists and animators

Source: Bureau of Labor Statistics, U.S. Department of Labor, Occupational Employment Statistics

Data on the locations of VC firms are from the MoneyTree™ Report, prepared by PricewaterhouseCoopers/National Venture Capital Association based on the survey data from

Thomson Financial[8]. The MoneyTree™ Report is a quarterly study of venture capital investment activity in the United States beginning with 1995. However, the industry-specific VC list is available only after the fourth quarter of 2001. This study identifies from this report 233 VCs that were actively invested in biopharmaceutical ventures and 712 VC firms that were actively investing in IT service ventures during the period from the fourth quarter of 2001 to the third quarter of 2002[9]. In 2005, the number of VC firms that were funding biopharmaceutical and IT service start-ups was 298 and 550, respectively. Some venture capital firms have established satellite offices. Unfortunately, the MoneyTree™ Report, like other commonly-used venture capital data sources (e.g., SDC Venture, see Stuart and Sorenson(2003)), only reports the location of the headquarters office for each VC firm. Therefore, our calculations of the distances of biotechnology firms and geographic areas to VCs will be biased downward for some metropolitan areas that only have VC branch offices.

The aggregate immigrant data for each MSA in the U.S. are from the American Communities Project, presented jointly by the Initiative in spatial Structures in the social sciences, Brown University, and the Lewis Mumford Center, University at Albany[10].

The academic R&D expenditure data was compiled from the National Science Foundation's (NSF's) Survey of Research and Development Expenditures at Universities and Colleges.

Data on industrial employment and establishments are mainly retrieved from the County Business Pattern (CBP) provided by the Census Bureau. The CBP is an annual series that reports subnational economic data by industry. Statistics include number of establishments, payroll

8 Money Tree Report, https://www.pwcmoneytree.com/MTPublic/ns/index.jsp
9 The 'biopharmaceutical' industry used in this study corresponds to the 'biotechnology' industry classification adopted in the MoneyTree™ Report. The IT Service industry in this study includes both "IT Service" and "Software" the MoneyTree™ Report.
10 Data source: http://www.s4.brown.edu/cen2000/data.html, accessed on 8/30/2007.

(annual and 1st quarter), and number of employees, etc. The CBP data has three major limitations that could affect the results of this study. First, the data is not completely comparable over time because of the change of industrial category system. Data for 1997 and earlier years are based on the SIC System. Since 1998, it has transferred to the NAICS system. Second, data for many counties are only a rough estimate. The CBP reported "flags' (A, B, C, etc.) to avoid disclosing information in states and counties where industry participants can be easily identified. This study substituted the "flags" by the midpoint of the range shown in Table 4. Third, I found data for some MSAs in certain years are extremely unbelievable. So, I check the corresponding values reported by the Quarterly Census of Employment and Salarys (QCEW) data which is provided by the Bureau of Labor Statistics. I then chose the values to be more consistent with the historical trend.

Table 4: Flags, Range, and Midpoints in the County Business Pattern Data

Flag	Range	Midpoint
A	0-19	10
B	20-99	60
C	100-249	175
E	250-499	375
F	500-999	750
G	1,000-2,4999	1750
H	25,00- 4,999	3750
I	5,000-9,999	7,500
J	10,000-24,9999	17,500
K	25,000-49,9999	37,500
L	50,0000-99,9999	75,000
M	100,0000 and more	100,000

This study uses the 1997 Input-Out (I-O) Benchmark account[11] published by the Bureau of Economic Analysis (BEA) of the U.S. Department of Commerce to determine the most important specialized buyer- and supplier- industries for the biopharmaceutical and IT service

[11] Source: http://www.bea.gov/industry/io_benchmark.htm#2002data, Bureau of Economic Analysis, U.S. Department of Commerce.

Industries. The 1997 Benchmark I-O account is the latest in a series of benchmark accounts that provide an extensive accounting of the commodity production and consumption in the U.S. economy[12]. We use its 'MAKE' table to identify the suppliers, and the 'USE' table to define the buyers. For example, to determine the suppliers of the biopharmaceutical industry, I first rank all the industries in the 'MAKE' table according to their sales to the biopharmaceutical industry. Industries with a rank of 10th or lower percentiles are deleted because of their trivial contribution to the biopharmaceutical industry. I then rank the rest industries according to the percentage of their output sold to the biopharmaceutical industry and select the top three industries as the most specialized and important suppliers of the biopharmaceutical industry. Similarly, I can use the 'USE' table to determine the buyers of the biopharmaceutical industry. Ketelhohn's (2002) has employed a similar approach to determine the buyer and supplier industries for the advertising, pharmaceutical, and semiconductor industries.

Industries in the 1997 Input-Out Benchmark tables were defined by a six-digit NAICS code. Specifically, the biopharmaceutical industry in this study corresponds to the industry code '325400' (Pharmaceutical and medicine manufacturing). The IT service industry encompasses four six-digit codes: 511200 (Software publishers), 541511(Custom computer programming services), 541512 (Computer systems design services), and 54151A (Other computer related services). Table 5 and 6 show the resulting set of top suppliers and buyers of the biopharmaceutical and IT service Industries.

[12] The preliminary 2002 benchmark Input-Out tables are available now. However, industries in these tables were defined only at a three-digit NAICS level, and thus don't provide information this study needs.

Table 5: Supplier- and Buyer- Industries of the Biopharmaceutical Industry

Supplier NAICS code	Supplier- industry name	Buyer NAICS code	Buyer-industry name
334516	Analytical laboratory instrument manufacturing	622	Hospitals
311514	Dry, condensed, and evaporated dairy products	621	Other ambulatory health care services
339112	Surgical and medical instrument manufacturing	54194	Veterinary services

Table 6: Supplier- and Buyer- Industries of the IT Service Industry

Supplier NAICS code	Supplier- industry name	Buyer NAICS code	Buyer-industry name
334113	Computer terminal manufacturing	334111	Electronic computer manufacturing
5612	Facilities support services	334113	Computer terminal manufacturing
5142	Data processing services	334112	Computer storage device manufacturing

The small firm birth rate data at metropolitan level is from 'Dynamic Firm Size' database prepared by the Office of Advocacy, U.S. Small Business Administration.

The industrial establishment data are from retrieved from the County Business Pattern data provided by the U.S. Census Bureau. The county-level land size data is Census 2000 Summary File 1 provided by the U.S. Census Bureau. The county-level land size data is then aggregated to the MSA level based upon the 2003 December MSA definition.

3.3. Models

This study employs both cross-sectional and two-period panel models to investigate the impacts of regional characteristics on average NTBF performance. A Zero-Inflated Negative Binomial (ZINB) model specification is used to account for the distributional characteristics of the count data which will be explained in next section. To check the robustness of the results, the same regression models will also be applied to a set of distance-weighted variables and a larger size of sample (include both the leading and catch-up regions).

3.3.1. Cross-sectional estimation

The cross-sectional model investigates the causal relationship between a catch-up region's initial conditions and its subsequent NTBF performance. The dependent variable IPO_Target_{im} is the total number of NTBF IPO and M&A events occurred in industry i within a MSA m during the period from 1996 to 2005. The independent variables include measures for local scientists job market conditions, venture capital, cultural diversity, university research expenditure, industrial structures, and entrepreneurial climate in 1995, one year prior to our study period. However, due to the data availability, the measures for some independent variables were in 1999 or later year (see table 2 for details). Given that most regional variables don't change dramatically over a very short period of time, such an approach is not expected to bias the findings substantially. The control variables include the industrial employment share, establishment density, number of small firms, and total employment in a metro area. The sample size for the biopharmaceutical industry is 168, and 301 for the IT service industry. Mathematically, the full cross-sectional model can be expressed as:

Equation (1):

$$\begin{aligned} IPO\_Target_m = f\,(&\beta_0 + \beta_1 Job\_share_m + \beta_2 salary\_log_m + \beta_3 Relative\_salary_m + \beta_4 VC\_log_m \\ &+ \beta_5 VC\_log\_sq_m + \beta_6 Immi\_share_m + \beta_7 Univ\_share_m + \beta_8 Ind\_lq_m \\ &+ \beta_9 Buyer\_lq_m + \beta_{10} Supplier\_lq_m + \beta_{11} Firm\_birth\_rate_m + \beta_{12} Est\_density_m \\ &+ \beta_{13} Ind\_SME\_log_m + \beta_{14} Ind\_SME\_log\_sq_m \\ &+ \beta_{15} Employment_m + \varepsilon\,) \end{aligned}$$

.

3.3.2. Two-period panel data estimation

The panel data model investigates the temporal stability of the causal relationships between location and NTBF performance. Although the region-level factors may change dramatically at the societal level over periods greater than 10 years, they change very little over short periods of time(Gartner and Shane 1995). Therefore, in this study, I partition the 10-year-long study period only into two five-year sub periods. The sample size is 168 for the biopharmaceutical industry, and 301 for the IT service industry. Due to the small sample size, this two-period panel data model doesn't account for the group fixed effect. The occurrence of IPO and M&A events is very sensitive to the overall capital market condition, which fluctuates over time. Thus, the panel data model includes one time dummy variable, period_dummy, to account for the time fixed effect. The dependent variable is number of NTBF IPO and M&A events that occurred in industry I within MSA m over 1996-2005 or 2001-2005. Most independent and controlled variables were measured in 1995 and 2000. For variables that were first measured in 1999 or later year, their subsequent values were measured five years later. This guaranteed that all variables were updated over the same length of time gap. Mathematically, the full panel data model can be expressed as:

Equation (2):

$$\begin{aligned}\text{IPO\_Target}_{imt} = f\,(&\beta_0 + \beta_1\text{Job\_ share}_{mt} + \beta_2\text{salary\_ log}_{mt} + \beta_3\text{Relative\_ salary}_{mt} + \beta_4\text{VC\_ log}_{mt}\\ &+ \beta_5\text{VC\_ log\_sq}_{mt} + \beta_6\text{Immi\_ share}_{mt} + \beta_7\text{Univ\_ share}_{mt} + \beta_8\text{Ind\_ lq}_{mt}\\ &+ \beta_9\,\text{Buyer\_ lq}_{mt} + \beta_{10}\text{Supplier\_ lq}_{mt} + \beta_{11}\text{Firm\_birth\_ rate}_{mt}\\ &+ \beta_{12}\text{Est\_ density}_{mt} + \beta_{13}\text{Ind\_SME\_log}_{mt} + \beta_{14}\text{Ind\_ SME\_log\_sq}_{mt}\\ &+ \beta_{15}\text{Employment}_{mt} ++ \beta_{16}\text{Period\_dummy} + \varepsilon_{mt}\,)\end{aligned}$$

3.3.3. ZINB model specification

As discussed above, the dependent variable in this study is the number of biopharmaceutical or IT service IPO and M&A events occurred in a region within a specific time period. This is an event count variable which has limited non-negative integer values and is highly skewed towards the lower values. When the ordinary least squares (OLS) method is applied to such event count data, it usually results in biased, inefficient, and inconsistent estimates (Long 1997). Thus, it is much safer to use models specifically designed for count outcomes in this study.

The ZINB model is chosen mainly for two reasons. First, the Negative Binomial Regression Model (NBRM) is preferred to the Poisson Regression Model (PRM) because overdisperson is apparent in the data. The PRM assumes that observed count is drawn from a Poisson distribution and that the equality of its conditional mean and conditional variance holds. When Var(y|x) > E(y|x), we are said to have overdispersion. Table 7 presents the preliminary evidence of overdispersion in our data. That is, the observed unconditional variance is much greater than the unconditional mean in both the cross-sectional and panel data models for two industries. The likelihood ratio tests for overdispersion (examines the null hypothesis of alpha=0), which will be reported in the Chapter 4 and 5, provide further evidence of overdispersion. In the presence of overdispersion, the standard errors in the PRM will be biased downward, resulting in spuriously large z-values and spuriously small p-values (Long & Freese, 2001, p. 243). The NBRM is a standard method used to model overdispersed Poisson data. The NBRM estimation incorporates observed and unobserved heterogeneity into the conditional mean, mu=exp(xb+e) (Long & Freese, 2001). Thus, the conditional variance of y becomes larger than its conditional mean, E(y|x)=mu, which remains unchanged.

Table 7: Descriptive Statistics of the Dependent Variable

Industry		N	Mean	Std Dev	Min.	Max.
Biopharmaceutical	cross-sectional model	160	1.381	3.003	0	26
	panel model	320	0.691	1.751	0	19
IT Service	cross-sectional model	301	7.970	19.072	0	189
	panel model	602	3.985	9.689	0	108

Second, the zero-inflated version of NBRM is favored over the standard NBRM due to the issue of excess zeros. Because this study focuses on the catch-up regions, the dependent variable has many zeros both in the cross-sectional and panel data models. Table 8 reports that in the biopharmaceutical industry, almost 60 percent (95 of 160) of catch-up MSAs had 0 IPO and M&A events between 1996 and 2005. The IT service industry had a smaller proportion of zero cases, but it is still over 33 percent (102 of 301). The excess-zeros situation in the two-period panel models is no better than that of the cross-sectional models. The Zero-inflated version of Negative Binomial model respond to the failure of standard NBRM models to account for excess zero. The vuong test results, which will be reported in Chapter 4 and 5, show that the ZINB is favored over NBRM in this study.

Table 8: Number of Zero and Non-Zero Observations for the Dependent Variable

Industry	Cross-sectional model			Two-period panel model		
	Total	Zeros	Non-zero	Total	Zeros	Non-zero
Biopharmaceutical	160	95	65	320	231	89
IT Service	301	102	199	602	283	319

The ZINB handles overdispersion by changing the mean structure to explicitly model the production of zero counts (Long & Freese, 2001). The model assumes that there are two latent groups: always-0 group that has an outcome of 0 with a probability of 1; not-always-0 group that might have a zero count, but there is a nonzero probability that the observation has a positive count. Since there is no observed variable indicating group membership, we do not know whether an observation is in the Always-0 group or Not-Always-0 group (Long & Freese, 2001, p. 251).

The cross-sectional version of ZINB model is implemented in the following three steps:

Step 1) Use the logit model to predict the probability of being Always-0 group for each metro area i. Because there is no theory suggesting which variables are the best predictors of the probability of being always-0 group, I use all the independent variables plus two control variables (est_density and ind_SME_log) as the predictors[13]. Here these variables are referred to as inflation variables since they serve to inflate the number of 0s. Suppose y_{im} = equation (1) and the probability of being in Always-0-group for individual MSA m is ψ_{im}, we have:

$$\varphi_{im} = \frac{\exp(y_{im})}{1+\exp(y_{im})} \qquad \text{Equation (3)}$$

Step 2) Model the probability of each count (including zeros) for those in Not-Always-0 group;

Pr (IPO_target $_{im}$ | x_i , being 'not-always-0' group)

$$= \frac{\Gamma(y_i + \alpha^{-1})}{y_i!\Gamma(\alpha^{-1})} + \left(\frac{\alpha^{-1}}{\alpha^{-1}+\mu_i}\right)^{\alpha^{-1}} + \left(\frac{\mu_i}{\alpha^{-1}+\mu_i}\right)^{y^i}$$

$$\Pr(IPO\_t\arg et_{im} \mid x_{im}, being\ 'not\ alway - 0' group\ )$$

[13] In some model specifications defined in Chapter 4 and 5, the algorithm doesn't converge to a solution if I use all the independent variables as the inflation variables. If this happened, I first drop the squared terms of independent variables, and then try to drop one or two variables if it still doesn't converge. The two control variables (ind_share and employment) are excluded because they are highly correlated with Ind_SME_lg in both industries.

$$= \frac{\exp(y_{im})}{1+\exp(y_{im})}$$

Step 3) Compute observed probabilities as a mixture of the probabilities for the two groups.

The panel data version of ZINB model can be implemented in the similar way except that the index is 'imt' instead of 'im'.

3.3.4. Robustness check methods

This study employed three approaches to check the robustness of the results we will get from the above model specification.

First, conduct Breusch-Pagan / Cook-Weisberg heteroskedasticity test for each model specification. If the 'constant variance of error term' assumption is violated, run the regression models with robust option. In stata, when we use the' robust' option, it specifies that the Huber/White/sandwich estimator of variance be used in place of the traditional calculation.

Second, estimate the full model by including both the leading and catch-up regions. By doing so, I can explore which causal relationships are unique to the catch-up regions.

Third, estimate the cross-sectional and panel full model based upon a set of distance-weighted variables. Stuart and Sorenson (2003) point out that the traditional within-region measures ignore the presence of many objects that fall beyond the arbitrary geographic boundary, even if they lie very near to the borderline. For example, when we measure the impacts of venture capital firms on firm performance in a MSA, we should not only account for venture firms that are located within that MSA, but also for those located in immediately neighboring MSAs. This is particularly crucial when the spatial unit of study is at the county, ZIP code, or smaller level. This study is based upon the metropolitan area level, so the 'neighboring' effect is unlikely to be a serious problem. However, some MSAs are very close to each other so that the 'neighboring' effect can not be simply ignored. Therefore, as a robustness check, this study will replicate the full models based upon a set of distance-weighted measures. The distance-weighted measures are constructed in the following way:

Step 1, identify the most 'crowded' zip code in each MSA based upon 2005 5-digit ZIP Code Business Pattern data provided by the Census Bureau. The most 'crowded' zip code refers to the area that hosts the largest number of establishments in the industry. Thus, the corresponding ZIP code of the same MSA could vary across industries. For example, for the Atlanta-Sandy Springs-Marietta (GA) metropolitan area, the most crowded ZIP code in the biopharmaceutical industry is '30071'. In the IT service industry, however, the most crowded ZIP code is '30004'. By focusing on the 5-digit ZIP code level, this study aims to create a finer distance measure of economic exchange between two MSAs.

Step 2, obtain the longitude and latitude at the center of the most crowded 'ZIP' code in each MSA, and follow the formula used by Stuart and Sorenson (2003, p.238) to calculate the physical distance between the two points, i and j, as

$$d_{ij} = 3437*\{\arccos [\sin(lat_i)*\sin(lat_j)+\cos(lat_i)\cos(lat_j)\cos(|long_i - long_j|)]\}$$

where latitude (lat) and longitude (long) are measured in radians. The constant 3437 converts the distance into units of miles. The distances among all the leading and catch-up MSAs are calculated. That is, in the end, I obtain a 168*168 dimensions of distance matrix for the biopharmaceutical industry, and a 316*316 dimensions of distance matrix for the IT service industry.

Step 3, weight the contribution of each MSA to the focal MSA by the inverse of their distance, and then sum these weighted contributions across all MSAs to yield a distance-weighted value for the focal MSA. For example, the value of distance-weighted life science academic research for MSA i is calculated as:

$$\mathrm{Univ\_share\_weighted}_{\mathrm{i}} = \sum_{\mathrm{j}=1}^{168} \frac{\mathrm{Univ\_lf}_{\mathrm{j}}}{1 + \mathrm{d}_{\mathrm{ij}}} \quad \text{when } \mathrm{i} = \mathrm{j}, \mathrm{d}_{\mathrm{ij}} = 0$$

Similarly, the value of distance-weighted computer science academic research for MSA i is calculated as:

$$Univ\_share\_weighted_i = \sum_{j=1}^{316} \frac{Univ\_cs_j}{1 + d_{ij}} \quad when\ i = j, d_{ij} = 0$$

Among the independent and control variables defined in section 3.2, the following five variables are most likely to have spill over effects on the adjacent regions: scientist job market size, venture capital firms, immigrants, academic R&D expenditure, and industrial establishments. Correspondingly, I follow the above approach to recalculate the values of Job_share, VC_log, VC_log_sq, Immi_share, Univ_share, Ind_sme_log, and Ind_sme_log_sq. I then re-estimate the cross-sectional and two-period panel models using these distance-weighted variables as a robustness check.

CHAPTER 4: BIOPHARMACEUTICAL INDUSTRY

This chapter analyzes the spatial heterogeneity in the performance of US biopharmaceutical NTBFs during the period from 1996 to 2005. The basic industry characteristics are discussed in Section 4.1. Section 4.2 describes the spatial distribution of NTBF IPO and M&A activities in the industry. Section 4.3 uses a set of Zero Inflated Negative Binomial regression models to investigate the determinants of the average NTBF performance in a region. Section 4.4 summarizes the main findings.

4.1. The Biopharmaceutical Industry

This study defines the biopharmaceutical industry by SIC 283 prior to 1997 and NAICS 32541 thereafter[14]. Some authors (Guo, Lev et al. 2005) define the modern 'biotechnology' industry in a similar way except that they also include the sector of 'commercial physical and biological research' (SIC 8731 prior to 1997 and NAICS 541712 thereafter). This study is limited to the manufacturing firms in the industry, mainly because data for biological research is not separated from the total physical and biological research after 1997. According to the NAICS classification, the biopharmaceutical firms primarily engaged in one or more of the following activities: "(1) manufacturing biological and medicinal products; (2) processing (i.e., grading, grinding, and milling) botanical drugs and herbs; (3) isolating active medicinal principals from botanical drugs and herbs; and (4) manufacturing pharmaceutical products intended for internal

[14] Some authors use similar SIC or NAICS codes to define the 'biotechnology' industry. It corresponds to two SIC code: 283, and 8731. Or, it covers two 2002_NAICS codes: 3254-Pharmaceutical and medicine manufacturing; and 54171-Scientific R&D services

and external consumption in such forms as ampoules, tablets, capsules, vials, ointments, powders, solutions, and suspensions”[15].

This section characterizes the US biopharmaceutical industry as an industry with a relatively small size, plenty of high-paid jobs, high barrier for entry, and skewed spatial distribution.

4.1.1. Industry employment and earnings

The biopharmaceutical industry defined in this study includes both traditional large pharmaceutical producers and new biotech start-ups. While the beginnings of the U.S. traditional pharmaceutical industry can be traced back to the 1820s (Ketelhohn, 2002), the modern biotechnology industry originated in the 1970s when the recombinant DNA technology was discovered by Stanford geneticist Stanley Cohen and University of California, San Francisco, biochemist Herbert Boyer (Zhang and Patel 2005).

Despite its popularity in the economic development policy-maker community, the biopharmaceutical industry still remains relatively small. According to data from County Business Patterns, the Bureau of Census, the number of biopharmaceutical establishments was 1836 in 2005 (Table 9). In the same year, its total number of employees was 247,847, only accounting for 0.2 percent of total national employment. As will be discussed in next chapter, the biopharmaceutical industry is one-sixth the size of the IT service industry in terms of employment..

Over the last decade, the biopharmaceutical industry experienced only a moderate growth in employment. From 1995 to 2005, the biopharmaceutical industry employment increased by 21 percent, with an annual growth rate of 2%. This growth rate is no faster than the growth in the

[15] Source: http://www.census.gov/epcd/naics02/def/NDEF325.HTM#N3254.

national total employment. As a result, its share of total national employment was fairly stable over the last decade, with a range from 0.19 percent in 1999 to 0.22 in 2003. However, the biotechnology industry organization reports that biotech revenues grew from $12.7 billion in 1995 to $50.7 billion in 2005, with an annual growth rate of nearly 30 percent[16]. Also, Hecker (2005) projected that there is extensive room for this industry to grow in the near future. His study estimates that the biopharmaceutical industry will add 68,000, or 23 percent new jobs from 2002-12 (p.59). This is contrast to many High-Tech manufacturing industries in which employments are projected to decline substantially.

The biopharmaceutical industry generates high-paid jobs. The average annual salary of U.S. biopharmaceutical workers was $87,818 in 2005, $47,000 greater than the national average annual salary. Over the last decade, the ratio of industry to national average annual salary was consistently high, ranging from 1.9 in 1996 to 2.2 in 2005. Given its growth potential, the biopharmaceutical industry is likely to continue to be a key source of high-paid jobs in the future.

Table 9: U.S. Biopharmaceutical Industry Statistics: 1995-2005

Year	Establishments	Employment	Share of total US employment (Percent)	Annual average salary	Ratio of industry to national annual average salary
1995	1,529	204,851	0.204	51,610	1.82
1996	1,637	207,295	0.203	54,959	1.90
1997	1,704	212,610	0.202	59,921	1.94
1998	1,812	217,111	0.201	64,781	2.03
1999	1,830	218,804	0.198	69,100	2.06
2000	1,823	227,461	0.199	75,250	2.13
2001	1,825	233,503	0.203	72,394	2.00
2002	1,779	237,905	0.212	73,162	1.99
2003	1,825	251,855	0.222	78,036	2.07
2004	1,833	246,297	0.214	82,055	2.09
2005	1,836	247,847	0.213	87,818	2.16

[16] The Biotechnology Industry Organization, http://www.bio.org/speeches/pubs/er/statistics.asp, accessed on 9/4/2007.

Source: (1) Data for employment were from the County Business Pattern, Bureau of the Census, Department of Commerce; (2) Data for average salary were from Quarterly Census of Employment and Salarys (QCEW), Bureau of Labor Statistics, Department of Labor.

4.1.2. Industry structure

The biotech industry is associated with a relatively concentrated size structure as compared to the national average. Table 10 reports that in 2005, firms with no more than 99 employees accounted for 75.9 percent of total industry establishments. The percentage with 100-499 employees was 18.2. Firms with 500 employees or more accounted for 5.4 percent. Compared to the national average level, the biopharmaceutical industry has higher proportion of large companies and lower proportion of small firms. Such a pattern has been fairly stable during the past decade. Because barriers to entry in the concentrated industries are typically higher than in the slightly concentrated industry, it is expected that it is more difficult for catch-up regions to catch up in the biopharmaceutical industry than in other lower concentration industries like IT service industry.

Table 10: Share of Total Establishments by Firm Size Classes for Biopharmaceutical Industry and National Average: 1995-2005

Year	Firm with 1~99 Employees		Firm with 100~499 employees		Firm with 500 or more employees	
	Biopharmaceutical	National	Biopharmaceutical	National	Biopharmaceutical	National
1995	77.44	97.72	16.15	2.03	6.41	0.25
1996	78.56	97.70	15.64	2.05	5.80	0.25
1997	79.11	97.68	15.43	2.07	5.46	0.25
1998	80.02	97.63	15.29	2.12	4.69	0.26
1999	79.51	97.57	15.79	2.17	4.70	0.26
2000	79.21	97.51	15.96	2.22	4.83	0.27
2001	78.63	97.51	16.55	2.21	4.82	0.27
2002	77.74	97.68	17.14	2.07	5.12	0.25
2003	76.71	97.67	17.86	2.09	5.42	0.25
2004	76.21	97.67	18.22	2.08	5.56	0.25
2005	75.93	97.66	18.68	2.09	5.39	0.25

Source: the County Business Pattern, Bureau of the Census, Department of Commerce.

4.1.3. Spatial distribution

The U.S. biopharmaceutical industry is highly geographically concentrated. The traditional pharmaceutical firms were concentrated in the east coast and in a few areas around Lake Michigan. For example, the New York/New Jersey area was the home base for a number of large pharmaceutical companies like Johnson & Johnson, Pfizer, Merck & Co, Bristol-Myers Co., Squibb and Sons, Wyeth, etc. The Abbott Laboratories was located in Chicago, Illinois, and Eli Lilly & Co. was in headquartered in Indianapolis, Indiana. All these traditional pharmaceutical companies were founded in the 1800s or the beginning of the 19th century. Beginning in the early 1970s, the industry was heavily affected by the biotechnology revolution, which is based on the recombinant DNA technology. This round of biotechnology revolution was led by small entrepreneurial firms that tended to locate near the academic centers that created the relevant knowledge and the source of venture capital (Galambos and Sturchio 1998, p.254). As a result, new industrial clusters were mainly formed in the East and the South (Ketelhohn, 2002).

The present spatial distribution of U.S. biopharmaceutical industry reflects such a historical transformation. Appendix 2 presents the industry's geographical distribution at the state level between 1995 and 2004. In 2004, 32 of the 51 states (including the District of Columbia) had 1000 or more employees in this industry, accounting for 97.6 percent of total biopharmaceutical employment. California had the largest share of 15.1 percent, followed by New Jersey and New York, with shares of 14.0 percent and 8.7 percent, respectively. Over the last decade, 28 states had higher shares in 2004 than in 1995, and 15 states had lower shares. However, the magnitude of increase or decrease is only moderate for most states. Florida had experienced the largest share increase from 1.6 percent in 1995 to 2.6 percent in 2004, gaining

3356 jobs. Illinois had the largest share drop from 9.2 percent to 7.7 percent, with a loss of 1.5 share point, or equivalent to a job loss of 899. Moreover, nine of the ten largest states in terms of biopharmaceutical employment in 1995 continued to be ranked among top ten in 2004. Only Texas slid from the tenth in 1995 to the twelfth in 2004. The combined share of these ten states was 73.4 percent in 1995 and 71.8 percent in 2004.

At the Metropolitan area level, 240 of 361 MSAs had the presence of biopharmaceutical establishments in 2004. However, only 50 MSAs have more than 1000 employees and 26 MSAs have a size between 500 and 1000. Table 11 lists the twenty largest Metropolitan Areas in terms of industry employment share in 2004. New York, Chicago, and Philadelphia are the top three metropolitan areas, with share of 20.6 percent, 7.1 percent, and 5.1 percent, respectively. These three metropolitan areas accounted for 32.8 percent of total industry employment. They mainly represent the geographical locations of traditional pharmaceutical producers. It depicts a clearly defined biopharmaceutical corridor that starts in Philadelphia and ends in Boston. Four of ten largest metropolitan areas are in California: San Francisco, Los Angeles, Oxnard, and San Diego.

In terms of the changes in industry share from 1995 to 2004, 80 of the 168 metropolitan areas that have data available in 1995, 2000, and 2004 had higher shares in 2004 than in 1995, 67 had lower shares, and 35 had the same portion of industry employment[17]. Boston-Cambridge-Quincy (MA-NH) had experienced the largest share increase, followed by San Francisco-Oakland-Fremont and Oxnard-Thousand Oaks-Ventura. All of these three metropolitan areas underwent increases of more than one share point, or equivalent of more than 3000 employees. In contrast, New York-Northern New Jersey-Long Island, Chicago-Naperville-Joliet, Rochester,

[17] CBP data is not completely comparable over time.

NY, and San Jose-Sunnyvale-Santa Clara, CA had the largest share drop, ranging from -3.1 to -1.1 share point.

Table 11: Twenty Largest Biopharmaceutical Metropolitan Areas in 1995 and 2004

Metropolitan Statistical Areas	2004 Share	2004 Rank	2000 Share	2000 Rank	1995 Share	1995 Rank
New York-Northern New Jersey-Long Island, NY-NJ-PA	20.58	1	21.61	1	23.67	1
Chicago-Naperville-Joliet, IL-IN-WI	7.68	2	8.17	2	9.44	2
Philadelphia-Camden-Wilmington, PA-NJ-DE-MD	5.11	3	4.60	4	4.30	4
San Francisco-Oakland-Fremont, CA	5.01	4	3.43	5	3.85	5
Los Angeles-Long Beach-Santa Ana, CA	4.83	5	5.85	3	5.53	3
Boston-Cambridge-Quincy, MA-NH	3.42	6	2.68	7	2.25	6
Oxnard-Thousand Oaks-Ventura, CA	2.93	7	2.33	8	1.83	13
Indianapolis, IN	2.44	8	3.30	6	1.83	11
San Diego-Carlsbad-San Marcos, CA	1.74	9	1.52	13	1.43	16
Miami-Fort Lauderdale-Miami Beach, FL	1.70	10	0.89	22	0.74	34
Kalamazoo-Portage, MI	1.53	11	1.71	9	1.91	9
Dallas-Fort Worth-Arlington, TX	1.46	12	1.44	15	2.06	8
Norwich-New London, CT	1.44	13	1.47	14	1.24	18
St. Louis, MO-IL	1.41	14	1.24	17	2.10	7
Baltimore-Towson, MD	1.33	15	1.12	18	1.38	17
Raleigh-Cary, NC	1.32	16	1.53	12	1.67	15
New Haven-Milford, CT	1.20	17	1.64	10	1.83	12
Kansas City, MO-KS	1.07	18	1.31	16	1.00	20
Tampa-St. Petersburg-Clearwater, FL	0.92	19	0.94	19	0.62	38
Bridgeport-Stamford-Norwalk, CT	0.91	20	0.76	28	0.85	23

Source: the County Business Pattern, Bureau of the Census, Department of Commerce.

As discussed in Chapter 3, this study defines a MSA as a ‘leading’ region if its high tech index value in 1995 was ranked 95th percentile or above. Such an approach results in 8 leading regions and 160 catch-up regions in the biopharmaceutical industry. Table 12 reports that in 1995, the 160 catch-up MSAs accounted for 47.5 percent of total biopharmaceutical employment, 63.18 percent of total biopharmaceutical establishments, and 59.07 percent of small and medium biopharmaceutical establishments. From 1995 to 2000, the catch-up regions haven’t gained much share in all these three measures.

Table 12: Share of Biopharmaceutical Industry for Leading and Catch-Up Regions,

in 1995 and 2000

	1995			2000		
	Emp. Share	Estab. Share	SME Estab. Share	Emp. Share	Estab. Share	SME Estab. Share
Leading regions	52.48	36.82	40.93	49.73	34.83	40.87
Catch-up regions	47.52	63.18	59.07	50.27	65.17%	59.13

Source: the County Business Pattern, Bureau of the Census, Department of Commerce.

4.2. IPO and M&A Activities in the Biopharmaceutical Industry

Of the 173 biopharmaceutical firms launched their IPO during the period from 1996 to 2005, 149 had fewer than 500 employees and were less than 20-year old when they became public, and therefore meet our definition of NTBFs. Figure 4 shows the number of total and sample biopharmaceutical IPOs each year. It clearly depicts that the IPO market for the biopharmaceutical industry fluctuated over time. There were three relatively ‘hot’ periods which were associated with higher volume of issues and higher valuation by the public investors: 1996-1997, 2000, and 2004-2005. The rest years were relative ‘cold’ markets in which only a few companies went public. Overall, there were more IPOs in the first five years than in the second five years. Hence, when we analyze how the number of IPOs in a region varies over time, the industry-specific equity market condition should be controlled for.

The SDC database doesn’t keep employment and operating history data for many M&A deals. This makes it very difficulty to judge whether the acquired firm is an NTBF or not. For those missing employment data, this analysis only counts privately-owned acquired firms. This is because in general, privately-owned firms are more likely to be a small and young NTBF than the publicly-traded firms. During the study period, there were 534 privately-held biopharmaceutical firms which were merged or acquired. 79 of them had no location information. The remaining 455 deals were treated as NTBF M&A events. Figure 4 shows the number of missing cases is roughly proportional to the total number of acquisition activities each year. Therefore, it may not cause serious bias over time.

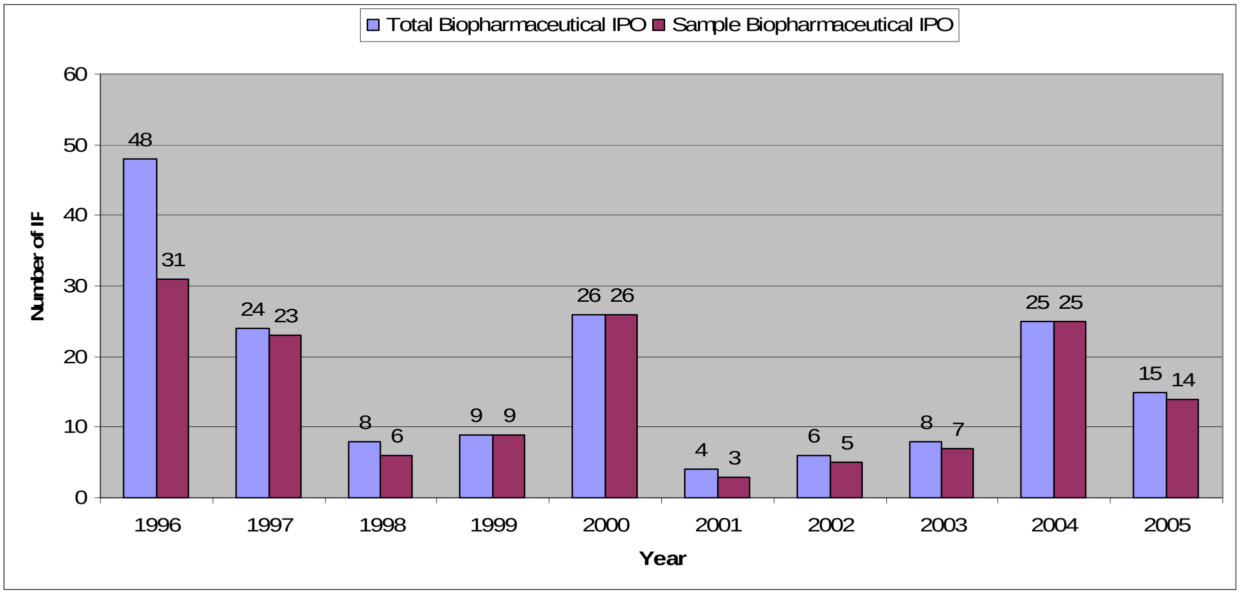


Figure 4: Number of Total and Sample Biopharmaceutical IPOs, 1996- 2005

Source: Securities Data Company, New Issues Database; Hoover's IPO list; Jay Ritter's 1975-2005 IPO dataset; and University of Chicago's Center for Research in Security Prices database.

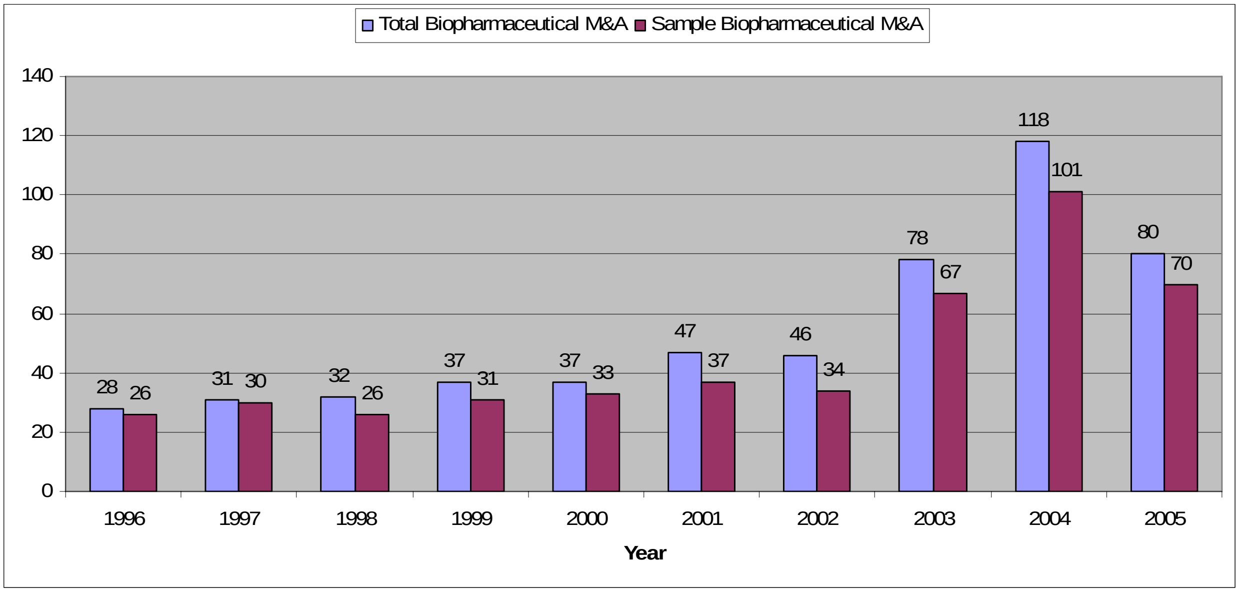


Figure 5: Number of Total and Sample Biopharmaceutical M&A, 1996- 2005

Source: Securities Data Company, New Issues Database.

The spatial distribution of the biopharmaceutical IPO and M&A events is highly skewed. Table 13 reports that 354, or 61.57 percent of total events occurred in the 8 leading MSAs. The catch-up regions, which comprise 160 MSAs, only had 221, or 38.43 percent. The share of IPO and M&A events for the catch-up regions is even lower than its share of industrial employment, which was 47.52 percent in 1995. The top five regions in terms of number of biopharmaceutical IPO and M&A events s are: New York-Northern New Jersey-Long Island (NY-NJ-PA), Boston-Cambridge-Quincy (MA-NH), San Diego-Carlsbad-San Marcos(CA), San Francisco-Oakland-Fremont (CA), and San Jose-Sunnyvale-Santa Clara (CA). These five regions in total had 290, or 50.43 percent of total deals. 73 of 168 MSAs had one or more deals, and the remaining 95 MSAs had no deal at all. Among the 160 catch-up regions, only 65 MSAs had one or more deals. Washington-Arlington-Alexandria (DC-VA-MD-WV) was ranked first, with 26 deals. This is followed by Atlanta-Sandy Springs-Marietta (GA), Miami-Fort Lauderdale-Miami Beach (FL), and Seattle-Tacoma-Bellevue (WA), with 10 deals each. Figure 3 depicts the spatial distribution of biopharmaceutical IPO and M&A events in catch-up Regions.

Table 13: Statistics of Biopharmaceutical IPO and M&A Events for the Leading and Catch-Up Regions

	Sum	Mean	Max	Min	Std
Leading regions	354	44.25	71	13	22.52
Catch-up regions	221	1.38	26	0	3.00

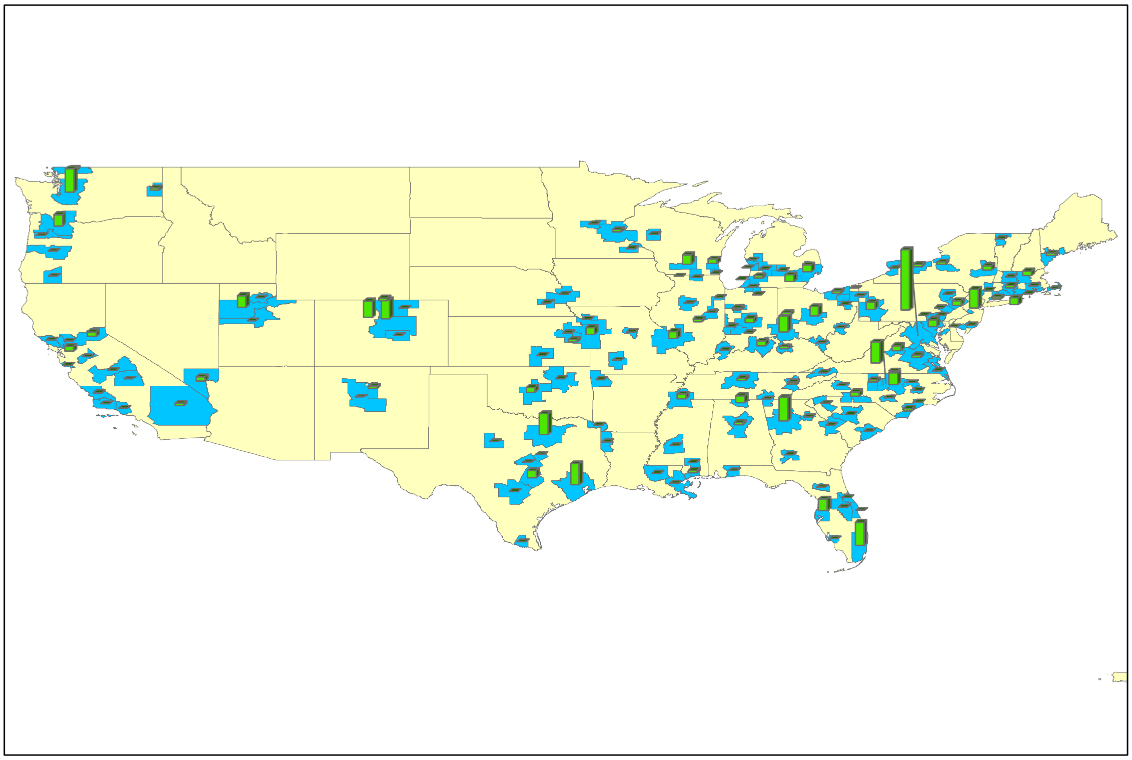


Figure 6: Spatial Distribution of Biopharmaceutical IPO and M&A Events in Catch-Up Regions

4.3. Determinants of Regional Variations in Biopharmaceutical NTBF performance

4.3.1. Models

Based on the research model and variable measures described in Chapter 3, this study employs both the cross-sectional and two-period panel ZINB models to identify the causal factors that affect the average biopharmaceutical NTBF performance in a catch-up region. The cross-section models predict the number of biopharmaceutical IPO and acquisition events that occurred in a region between 1996 and 2005 by a set of regional characteristics in 1995 or other years. These models are limited to 160 MSAs where the biopharmaceutical industry consistently existed from 1995 to 2004 so that all the observations had the same exposure time to the events. The two-period panel data models predict the number of biopharmaceutical IPO and acquisition events that occurred in a region during two sub periods (1996-2000 and 2001-2005). The values for the predictor variables were updated every five years. The sample size for the panel data model is 320.

The Breusch-Pagan / Cook-Weisberg heteroskedasticity test statistic is 631.18 ($p<0.00$) for the cross sectional model, and 1513.9 ($p<0.00$) for the two-period panel data, suggesting the presence of heteroskedasticity in the data. Therefore, this study only reports the robust estimation results.

Table 14 and 15 report the descriptive statistics and table 16 and 17 presents the Pearson correlation coefficient matrix both for the cross-sectional and panel data. Some correlation coefficients are very large, indicating that multicollinearity might be a concern in the data. For example, Table 16 reports that in the cross-sectional dataset, the control variable, a region's total job market size (emp_share), is strongly correlated with its life scientist job market size (job_share), with a correlation coefficient of 0.84. This control variable is also highly correlated with the number of small biopharmaceutical firms (SME_Biotech_lg), with a correlation coefficient of 0.68. In addition, the correlation coefficient between the absolute salary level of life scientists (Salary_log) and relative salary level (Relative_salary) is 0.72. In the panel data correlation matrix (Table 17), a region's total job market size is also highly correlated with the life scientist job market size, with a correlation coefficient of 0.84. The full regression model specified in Chapter 3 includes all these highly correlated variables and therefore the 'multicollinearity' becomes a concern. The presence of multicollinearity doesn't affect the overall predictive power or the statistical significance of the model. However, it can increase the standard errors of the regression coefficients and consequently underestimate their significant level.

Two approaches are employed to address the potential multicollinearity problem. First, the control variable, a region's total job market size, is excluded from the full model. This control variable is not only highly correlated with some independent variables, but also with another control variable (Ind_sme_log) that measures the size of industrial cluster in a region. Including the total job market size variable into the model might risk increasing the standard errors of its highly correlated variables, and consequently results in unreliable statistical inference.

Second, besides the full model, three reduced models are constructed to distinguish the total effect of two life scientists salary level variables on firm performance from their direct effect (when hold the other variables constant). Specifically, Reduced model 1 only includes the absolute salary level variable (Absolute_salary). Reduced model 2 only includes the relative salary level variable (Relative_salary). Reduced model 3 also only includes the relative salary variable, but it is based upon a smaller dataset in which 7 outlier regions were excluded. Regions with high absolute salary level are typically with a high relative salary level because they are highly correlated. However, in some cases, regions with low absolute salary level can also have high relative salary level because the denominator, the average salary level for all occupational jobs are very low. The presence of such outlier regions may completely distort the causal relationship between the life scientist relatively salary ratio and firm performance. To identify these outlier regions, I calculate the percentile ranks based upon the absolute and relative salary level of local life scientists, respectively. Seven regions were defined as the outliers because the difference in the percentile ranks for these regions is over 60. See appendix 3 for the results of these three reduced models.

All the model specifications use the same set of 'inflate' variables that predict the probability of being in the 'always-0-group' for a region. Because there are no solid theoretical arguments that point out the most significant variables for predicting the group membership, all the independent and control variables except their squared terms are used as the predictors. However, the algorithm for ZINB didn't converge to a solution until two control variables: number of small biopharmaceutical establishments and establishment density were dropped. Chapter 5 will discuss a similar problem when the ZINB full model is applied to the IT service industry data. This suggests one drawback of using a complicated model specification.

Table 14: Descriptive Statistics for Biopharmaceutical Cross-Sectional Variables (N = 160)

Variable	Mean	Std Dev	Minimum	Maximum
IPO_Target	1.38	3.00	0.00	26.00
Job_share	0.26	0.39	0.01	2.10
Absolute_salary	0.99	0.13	0.57	1.27
Relative_salary	1.19	0.13	0.85	1.57
VC_firm	0.44	1.19	0.00	7.00
VC_firm _sq	1.59	6.09	0.00	49.00
Immi_share	4.68	4.31	0.49	24.69
Univ_share	0.35	0.62	0.00	3.74
Ind_lq	1.20	2.45	0.01	16.32
Buyer_lq	0.03	0.02	0.00	0.12
Supplier_lq	1.38	2.00	0.00	19.91
Firm_Birth_Rate	12.04	2.08	7.84	18.80
Ind_SME	4.74	5.50	0.00	31.00
Ind_SME _sq	52.59	134.27	0.00	961.00
Est_density	0.00	0.00	0.00	0.02
Emp_share	0.45	0.55	0.05	3.00

Table 15: Descriptive Statistics for Biopharmaceutical Panel Data Variables (N =320)

Variable	Mean	Std Dev	Minimum	Maximum
IPO_Target	0.69	1.75	0.00	19.00
Job_share	0.29	0.41	0.01	2.58
Absolute_salary	0.99	0.13	0.57	1.41
Relative_Salary	1.25	0.15	0.85	1.88
VC_firm	0.49	1.23	0.00	7.00
VC_firm _sq	1.75	6.57	0.00	49.00
Immi_share	5.84	5.04	0.49	31.21
Univ_share	0.34	0.63	0.00	4.14
Ind_lq	1.23	2.38	0.01	16.32
Buyer_lq	0.41	0.46	0.00	4.20
Supplier_lq	1.17	2.21	0.00	26.07
Firm_Birth_Rate	1.17	2.10	7.80	21.30
Ind_SME	5.17	6.18	0.00	37.00
Ind_SME_sq	64.81	168.61	0.00	1369.00
Est_density	0.00	0.00	0.00	0.04
Emp_share	0.45	0.55	0.05	3.02
Year_dummy	0.50	0.05	0.00	1.00

Table 16: Correlation Coefficient Matrix for the Biopharmaceutical Cross-Sectional Data

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
(1) IPO_Target	1.00													
(2) Job_share	0.72	1.00												
	<.0001													
(3) Absolute_salary	0.32	0.30	1.00											
	<.0001	0.00												
(4) Relative_salary	0.05	0.05	**0.72**	1.00										
	0.50	0.56	<.0001											
(5) VC_firm	0.59	0.66	0.30	0.07	1.00									
	<.0001	<.0001	<.0001	0.38										
(6) Immi_share	0.23	0.16	0.31	0.10	0.04	1.00								
	0.00	0.04	<.0001	0.23	0.65									
(7) Univ_share	0.54	0.57	0.23	0.05	0.68	0.08	1.00							
	<.0001	<.0001	0.00	0.49	<.0001	0.31								
(8) Ind_lq	-0.07	-0.12	-0.01	0.01	-0.07	-0.12	-0.07	1.00						
	0.39	0.15	0.93	0.89	0.39	0.13	0.40							
(9) Buyer_lq	0.05	0.04	0.13	0.06	0.07	0.07	0.09	-0.05	1.00					
	0.56	0.65	0.10	0.45	0.36	0.39	0.25	0.51						
(10) Supplier_lq	-0.02	-0.01	0.09	0.05	0.03	0.08	0.09	-0.06	-0.02	1.00				
	0.76	0.88	0.25	0.49	0.72	0.33	0.27	0.46	0.82					
(11) Firm_Birth_Rate	0.24	0.18	-0.04	0.06	0.14	0.24	0.11	-0.17	-0.08	-0.05	1.00			
	0.00	0.02	0.61	0.48	0.08	0.00	0.17	0.03	0.33	0.51				
(12) Ind_ SME	0.54	0.72	0.22	-0.04	0.53	0.21	0.48	-0.05	0.05	-0.02	0.24	1.00		
	<.0001	<.0001	0.01	0.65	<.0001	0.01	<.0001	0.51	0.53	0.82	0.00			
(13) Est_density	0.27	0.22	0.25	-0.05	0.17	0.11	0.17	0.25	0.18	0.04	-0.13	0.37	1.00	
	0.00	0.00	0.00	0.52	0.03	0.17	0.03	0.00	0.02	0.60	0.09	<.0001		
(14) Employment	0.72	**0.84**	0.28	-0.01	0.58	0.23	0.52	-0.17	0.08	-0.06	0.19	**0.77**	0.13	1.00
	<.0001	<.0001	0.00	0.90	<.0001	0.00	<.0001	0.03	0.32	0.44	0.01	<.0001	0.09	

Note: a. The square terms are not included in order to save space; b. the first contains the correlation coefficient, and the p-values are in the second row
b. Number of observations: 160

.

Table 17: Correlation Coefficient Matrix for the Biopharmaceutical Panel Data

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
(1) IPO_Target	1.00													
(2) Job_share	0.67	1.00												
	<.0001													
(3) Absolute_salary	0.34	0.42	1.00											
	<.0001	<.0001												
(4) Relative_salary	0.05	0.11	0.58	1.00										
	0.36	0.05	<.0001											
(5) VC_firm	0.57	0.70	0.37	0.08	1.00									
	<.0001	<.0001	<.0001	0.15										
(6) Immi_share	0.26	0.22	0.29	0.12	0.10	1.00								
	<.0001	<.0001	<.0001	0.04	0.08									
(7) Univ_share	0.45	0.58	0.29	0.05	0.65	0.10	1.00							
	<.0001	<.0001	<.0001	0.37	<.0001	0.08								
(8) Ind_lq	-0.06	-0.12	-0.05	-0.02	-0.08	-0.11	-0.08	1.00						
	0.25	0.03	0.39	0.72	0.15	0.06	0.17							
(9) Buyer_lq	-0.02	0.00	-0.06	0.33	-0.03	0.07	-0.02	-0.01	1.00					
	0.72	0.96	0.30	<.0001	0.57	0.22	0.66	0.89						
(10) Supplier_lq	-0.04	-0.04	0.13	0.06	0.00	0.02	0.06	-0.03	-0.12	1.00				
	0.51	0.44	0.02	0.28	0.97	0.72	0.30	0.58	0.03					
(11)Firm_Birth_Rate	0.20	0.16	0.07	0.01	0.15	0.30	0.10	-0.16	-0.21	-0.05	1.00			
	0.00	0.00	0.23	0.92	0.01	<.0001	0.09	0.00	0.00	0.39				
(12) Ind_ SME	0.56	0.75	0.36	0.07	0.60	0.29	0.48	-0.06	-0.06	-0.03	0.25	1.00		
	<.0001	<.0001	<.0001	0.18	<.0001	<.0001	<.0001	0.27	0.32	0.59	<.0001			
(13) Est_density	0.30	0.19	0.25	-0.04	0.12	0.15	0.13	0.19	0.00	0.01	-0.08	0.33	1.00	
	<.0001	0.00	<.0001	0.47	0.03	0.01	0.02	0.00	0.96	0.84	0.18	<.0001		
(14) Employment	0.62	**0.84**	0.38	0.06	0.64	0.27	0.52	-0.18	-0.09	-0.07	0.20	**0.79**	0.11	1.00
	<.0001	<.0001	<.0001	0.31	<.0001	<.0001	<.0001	0.00	0.11	0.23	0.00	<.0001	0.06	

Note: a. The square terms are not included in order to save space; b. the first contains the correlation coefficient, and the p-values are in the second row
b. Number of observations: 320

4.3.2 Results

This section presents the results for the empirical estimation of the effects of regional characteristics on the average biopharmaceutical NTBF performance in a region, measured by the number of IPO and M&A events. I first report the estimation results from the cross-sectional models. The temporal stability test findings based upon two-period panel data models are followed. Finally, I present the robustness check results which are based upon distance-weighted dataset.

(1) Cross-sectional models

Table 18 presents the coefficient estimates of the cross-sectional zero-inflated negative models both for catch-up regions and all the regions. In each model, the top set of coefficients corresponds to the NBRM for those in the 'Not- Always-0-Group'. The lower set of coefficients, labeled 'inflate', corresponds to the binary logistic model predicting membership in the group that always has zero counts. Table 18 also reports the results of alpha and Vuong test in the end. The alpha test in each model is larger than 1.96, which provides strong evidence of overdispersion. In the presence of overdispersion, "estimates from the Poisson Regress Model are inefficient with standard errors that are biased downward, even if the model includes the correct variables" (Long & Freese, 1997, p.246). Therefore, the negative binomial regression model is preferred to the Poisson regression model in this data. The Vuong test compares the zero-inflated regression model specification with the standard negative binomial regression model. The Vuong test statistic in each model is larger than 1.96, indicating that the zero-inflated regression model specification is favored in this data.

The size of local life scientist job market is found to have a significant and positive impact on the number of biopharmaceutical IPO and M&A events that occurred in a catch-up region which belongs to the 'not-always-0-group'. The estimated coefficient of this variable is 0.493 ($p < 0.1$), suggesting that for one unit (i.e. one percentage point) change in an MSA's life scientist job market share, the expected count of biopharmaceutical IPO and M&A events changes by a factor of exp(0.493), or 1.64, holding all other variables constant. This finding indicates that in catch-up regions, the clustering of life scientist occupational jobs plays a significant role in fostering the growth of local biopharmaceutical NTBFs. When the leading regions are included into the sample, this variable has a smaller and positive coefficient (0.223), but its significance diminished. In terms of predicting the odds of being the 'always-0-group', the variable is negative and insignificant for catch-up regions, but negative and significant ($p<0.1$) in the full sample model. Therefore, the results provide some evidence that the local scientist job market size in a catch-up or leading region has positive impact on biopharmaceutical NTBF performance.

There is little evidence that local life scientist salary level has significant impacts on NTBF performance in catch-up regions. The life scientist absolute salary rate, which is measured by the ratio of local average salary rate to the national average rate, has positive but insignificant impacts on the rate of IPO and M&A events. The relative salary ratio, which is calculated by dividing the average salary of 'life scientists' jobs in a region by the average salary of all occupational jobs in that region, has negative and insignificant impact. Table 3.1 in Appendix 3 reports the results of three reduced models defined in the above section. Neither variable is statistically significant even in the reduced model. Therefore, multicolinearity is not the cause for the insignificance of these two variables. However, in the full sample model which consists of both the leading and catch-up regions, both variables are statistically significant, with a significant level of 0.01 and 0.1 respectively. In addition, in the full model, the absolute salary rate variable is negative and significant ($p<0.1$) in terms of predicting the odds of 'always-being-0". Therefore, the results indicate that life scientist salary level matters more in leading regions than in catch-up regions. In other words, it suggests that one advantage of the leading regions is that they can provide nationally competitive high-paid jobs for talented life scientists and technicians.

Proximity to biotech venture capital firms is found to have non-linear but insignificant effect on the rate of expected biopharmaceutical IPO and M&A events both in the catch-up region and full sample models. The number of biotech VC firm has a positive coefficient, and its squared term has a negative coefficient. Both variables are statistically insignificant. However, in terms of predicting the odds of being the 'always-0-group', the venture capital is negative and highly significant ($P<0.01$) both for the catch-up and full sample models. This suggests that the presence of additional biotech VC firms will reduce the odds of not having biopharmaceutical IPO and M&A events. Therefore, the results provide some evidence that is consistent with previous studies. For example, Shane and Stuart (2002) have shown that linkages with venture capitalists were more likely to lead to success, especially as it related to reaching an IPO stage.

Cultural diversity has positive but insignificant impact on the performance of NTBF performance. The coefficient of foreign born population share is positive and insignificant for catch-up regions, and negative and insignificant in the full sample model. Furthermore, in predicting the likelihood of having zero IPO and M&A event, this variable has a positive sign, even though it is not statistically significant. This indicates that in the sample, regions that have high proportion of foreign-born immigrants are more likely to experience zero IPO and M&A events in the biopharmaceutical industry. Therefore, the results provide little evidence supporting

Florida's (2002) argument that diverse place are more attractive to creative people and therefore have better NTBF performance than regions that have higher barriers for immigrants.

The results provide little evidence of the direct and significant impact of academic life science research on the performance of local small biopharmaceutical firms. The estimated coefficient of life science academic research and development expenditure is positive but insignificant both in the catch-up and full sample models. In terms of predicting the odds of having no biopharmaceutical IPO and M&A events, the variable is negative but insignificant. There are two possible explanations for the insignificance of academic research. The first explanation, which here we don't have data to prove, is that academic scientists tend to collaborate more frequently with large biopharmaceutical firms than with the smaller ones. Another explanation is that there are few high-quality universities or research institutes in the catch-up regions.

The coefficients of three industrial structures variables are not statistically significant. The concentration of biopharmaceutical industry and the concentration of its supplier industries have negative and insignificant coefficients in two models. The concentration of buyer industries is positive but insignificant. As 'Inflation' variables, they are also insignificant. This finding suggest that the extent of industrial specialization in a region is not a determinant factor of the rate of biopharmaceutical IPO and M&A events, nor the likelihood of being 'always-zero-group'.

The findings provide strong evidence that the overall entrepreneurial climate in a region has significant and positive impacts on the performance of biopharmaceutical NTBFs. The local entrepreneurial climate, measured by the small firm birth rate, is positive and statistically significant at 0.1 level. The model predicts that for one percentage point change in the firm birth rate, the expected count will change by a factor of exp (0.12), or 1.13.

The three control variables, which are used to measure the characteristics of local industrial clusters, are insignificant in the catch-up model. However, in the full sample model, the number of small biopharmaceutical firms is positive and significant at 0.05 level. Its squared term is negative and highly significant ($p<0.01$). This result suggests that the size of industrial cluster has positive but non-linear effect on NTBF performance. The model predicts when the size of impact on industrial clusters on firm performance reaches the maximum when there are about 68 small biopharmaceutical firms. This finding is consists with Folta et al (2006). They found that in the US biotech industry, the diseconomies of agglomeration dominate the economic benefits of agglomeration when clusters exceed about 65 firms.

Table 18: Results of Cross-Sectional ZINB Models of Biopharmaceutical IPO and M&A Events between 1996 and 2005

Independent variables	**Catch-up regions (N=160)**	**Leading and catch-up regions (N=168)**
Life scientists job market share	**0.493*** (0.294)	0.223 (0.222)
Life scientists absolute salary ratio	3.179 (2.324)	**4.936***** (1.780)
Life scientists relative salary ratio	-2.824 (3.064)	**-3.805*** (2.202)
Number of biotech VC firms	0.283 (0.227)	0.019 (0.088)
Square of number of biotech VC firms	-0.044 (0.047)	-0.000 (0.002)
Foreign born population share	0.039 (0.058)	-0.002 (0.022)
Life science academic research share	0.155 (0.371)	0.247 (0.154)
Biopharmaceutical industry location quotient	-0.007 (0.062)	-0.005 (0.052)
Buyer- industry location quotient	3.007 (6.009)	2.499 (4.296)
Supplier- industry location quotient	-0.059 (0.090)	-0.063 (0.065)
Small firm birth rate	**0.123*** (0.072)	**0.181***** (0.053)
Number of small biopharmaceutical firms	0.123 (0.105)	**0.044**** (0.022)
Square of number of small biopharmaceutical firms	-0.004 (0.003)	**-0.000***** (0.000)
Number of small biopharmaceutical firms per square mile	48.462 (34.200)	20.277 (35.340)
Constant	-2.240 (1.828)	**-2.772*** (1.502)

Table 18 (continued)

Inflation variables	**Catch-up regions (N=160)**	**Leading and catch-up regions (N=168)**
Life scientists job market share	-5.758 (3.683)	**-7.020*** (3.860)
Life scientists absolute salary	-19.551 (19.555)	**-15.851*** (9.184)
Life scientists relative salary	6.660 (10.948)	6.077 (7.219)
Number of biotech VC firms	**-15.633***** (4.676)	**-19.322***** (2.582)
Foreign born population share	0.349 (0.361)	0.260 (0.201)
Life Science academic research share	-4.126 (11.058)	-3.160 (4.078)
Biopharmaceutical industry location quotient	-0.026 (0.308)	-0.067 (0.130)
Buyer- industry location quotient	-2.383 (32.234)	-2.913 (21.158)
Supplier- industry location quotient	-0.426 (0.399)	-0.489 (0.394)
Small firm birth rate	-0.301 (0.247)	-0.274 (0.186)
Constant	15.207 (25.569)	12.760 (11.356)
Log likelihood	-155.958	-208.024
Overdispersion test (alpha =0)	12.25	123.84
Vuong test (ZINB vs. NBRM)	4.45	2.50
Number of nonzero observations	65	73
Number of zero observations	95	95

Note: Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

(2) Two-period panel data models

Table 19 reports the coefficient estimates of two-period panel zero-inflated negative models both for catch-up regions and the full sample. In both models, the time fixed effect is controlled for by using one dummy variable, which equals 0 for the period from 1996 to 2000, and 1 for the period from 2001 to 2005. The MSA-level regional fixed effects are not controlled for in these models due to the limited number of observations. The state-level variables are not appropriate either because some MSAs are cross several states. Table 19 also reports the results of overdispersion and Vuong tests. Each statistic value is larger than 1.96, indicating that the NBRM is preferred to PRM, and ZINB is preferred to NBRM in this data.

The two-period panel results for the majority of predictors are qualitatively similar to the findings in the cross-sectional models, indicating the temporal stability of the causal patterns discussed in the previous section. In particular, life scientist job market size and small firm birth rate, which are the only two statistically significant variables in the cross-sectional model, now have bigger coefficients and higher significant level. In terms of predicting effects on the rate of biopharmaceutical NTBF IPO and M&A events, the coefficient of local life scientist job market size has been increased from 0.493 to 0.854, and its significant level has changed from 0.1 to 0.05. Similarly, the coefficient of small firm birth rate has been increased from 0.123 to 0.161, and its significant level has changed from 0.1 to 0.05. Therefore, the two-period panel data provides further evidence that local life scientist job market size and entrepreneurial climate are the two key determinants of biopharmaceutical NTBF performance in catch-up regions.

The coefficients for other independent variables are not statistically significant, but most of them keep the same sign as in the cross-sectional model. For example, the following variables are found to have positive impacts on the rate of biopharmaceutical IPO and M&A events in the sample catch-up regions: life scientist absolute salary ratio, cultural diversity, life science academic research. The impact of relative salary ratio and biopharmaceutical industry specialization is negative. Proximity to biotech venture capital firms has positive but non-linear effects. Only the sign of buyer- and supplier-industry location quotient has been flipped over.

The control variable, number of small biopharmaceutical firms per square mile, is found to be positive and significant ($p<00.05$), suggesting that local industrial cluster density has positive impact on NTBF performance, holding other variables constant.

The time fixed effect is positive and significant ($p<0.1$) in both the catch-up and full sample models. This result is consistent with the evidence that overall, the second 5-year-period (2001-2005) has more biopharmaceutical IPO and M&A events than in the first 5-year-period (1996-2000).

In terms of predicting the odds of ‘being-always-0-group’, the absolute salary ratio of local life scientists is found to be negative and significant ($p<0.1$). The result suggests that regions with higher absolute salary level of life scientists are less likely to experience have zero number of IPO and M&A events.

Table 19: Results of Two-Period Panel ZINB Models of Biopharmaceutical IPO and M&A Events between 1996 and 2005

Independent variables	Catch-up regions (N=320)	Leading and catch-up regions (N=336)
Life scientists job market share	**0.854**** (0.360)	0.220 (0.148)
Life scientists absolute salary ratio	1.068 (2.613)	1.760 (1.252)
Life scientists relative salary ratio	-2.698 (1.732)	-0.791 (1.810)
Number of biotech VC firms	0.284 (0.316)	0.063 (0.043)
Square of number of biotech VC firms	-0.040 (0.051)	-0.001 (0.001)
Foreign born population share	0.006 (0.024)	-0.011 (0.018)
Life Science Academic research share	0.073 (0.192)	0.102 (0.157)
Biopharmaceutical industry location quotient	-0.019 (0.091)	-0.087 (0.090)
Buyer- industry location quotient	-0.502 (1.568)	-0.677 (0.538)
Supplier- industry location quotient	0.018 (0.110)	0.058 (0.081)
Small firm birth rate	**0.161**** (0.072)	**0.175***** (0.065)
Number of small biopharmaceutical firms	0.084 (0.090)	0.022 (0.018)
Square of number of small biopharmaceutical firms	-0.002 (0.002)	-0.000 (0.000)
Number of small biopharmaceutical firms per square mile	**52.747**** (23.615)	**46.521***** (15.482)
Time fixed effect	**1.079*** (0.654)	**0.894*** (0.511)
Constant	-1.464 (2.578)	**-3.437**** (1.417)

Table 19 (continued)

Inflation variables	Catch-up regions	Leading and catch-up regions
Life scientists job market share	-2.389 (1.853)	**-3.527***** (1.253)
Life scientists absolute salary	**-24.003*** (12.442)	**-10.746***** (4.058)
Life scientists relative salary	2.687 (18.884)	**6.746*** (4.054)
Number of biotech VC firms	-3.881 (5.553)	-1.353 (1.218)
Foreign born population share	0.060 (0.257)	0.046 (0.063)
Life Science academic research share	-10.815 (18.632)	-2.321 (2.200)
Biopharmaceutical industry location quotient	0.140 (0.630)	-0.080 (0.152)
Buyer- industry location quotient	0.402 (9.153)	-1.326 (1.293)
Supplier- industry location quotient	0.580 (1.132)	0.160 (0.205)
Small firm birth rate	0.170 (0.884)	-0.056 (0.175)
Constant	**17.967**** (9.067)	4.830 (4.095)
Log likelihood	-226.414	-302.885
Overdispersion test (alpha =0)	9.92	124.76
Vuong test (ZINB vs. NBRM)	2.39	3.06
Number of nonzero observations	89	105
Number of zero observations	231	231

Note: Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

(3) Distance-weighted results

This study replicates the cross-sectional and two-period models to a set of distance-weighted variables. The following variables are recalculated using the method defined in Chapter 3: life scientist job market share, number of biotech VC firms and its squared term, foreign born population share, life science academic research, number of small biopharmaceutical firms, and its squared term. These variables are likely to have impacts on the neighboring metropolitan areas. Table 20 presents the results. Again, the overdispersion and Vuong test statistics in each model are larger than 1.96, indicating that the ZINB specification is still the best choice.

The distance-weighted models provide further evidence of the significance of local entrepreneurial climate, measured by the small firm birth rate. In the cross-sectional model for catch-up regions, the small firm birth rate variable is the only significant ($p<0.05$) predictor of the expected event count. In the two-period panel model, this variable is more significant ($p<0.01$), and its coefficient has been increased from 0.14 to 0.18. This variable is also positive and highly significant ($p<0.01$) in the full sample models. Therefore, the result shows that local entrepreneurship is a key determinant of biopharmaceutical NTBF performance both in catch-up and leading regions.

The cross-sectional and panel models for catch-up regions also provide evidence supporting the following causal relationships: (1) the job market size of life scientist has positive and significant impact ($p<0.05$) on biopharmaceutical NTBF performance in a catch-up region; (2) the impact of relative salary level is negative and significant ($p<0.1$); (3) proximity to biotech venture capital firms and local life scientist absolute salary ratio decrease the odds of having zero IPO and M&A events; and (4) the industrial cluster density has positive and significant ($p<0.05$) impact over time.

When the leading regions are added to the dataset, more variables become significant. For example, in the cross-sectional model, life scientist absolute salary ratio has positive and highly significant ($p<0.01$) impact on the rate of events. The coefficient of relative salary ratio is negative and significant at 0.1 level. More interesting, life science academic research is positive and significant ($p<0.1$). This finding suggests that the impact of life science academic research on biopharmaceutical NTBF performance is significant only after we take into account the evidence of leading regions, and the impact from both the focal metropolitan area and it neighboring regions.

Table 20: Distance-Weighted Results of ZINB Models of Biopharmaceutical IPO and M&A Events between 1996 and 2005

	Catch-up regions		Leading and catch-up regions	
Independent variables	**Cross-Sectional (N=160)**	**Panel (N=320)**	**Cross-Sectional (N=168)**	**Panel (N=316)**
Life scientists job market share	0.255	**0.547****	0.118	0.108
	(0.220)	(0.252)	(0.165)	(0.106)
Life scientists absolute salary ratio	2.936	0.790	**4.817*****	1.758
	(1.999)	(2.435)	(1.656)	(1.170)
Life scientists relative salary ratio	-2.573	**-2.612***	**-3.580***	-0.569
	(3.855)	(1.502)	(2.151)	(1.617)
Number of biotech VC firms	0.211	0.219	-0.004	0.033
	(0.179)	(0.209)	(0.071)	(0.033)
Square of number of biotech VC firm	-0.033	-0.030	0.000	-0.000
	(0.047)	(0.035)	(0.002)	(0.001)
Foreign born population share	0.042	0.009	0.009	-0.002
	(0.086)	(0.021)	(0.017)	(0.014)
Life Science academic research share	0.141	0.075	**0.210***	0.097
	(0.399)	(0.132)	(0.110)	(0.109)
Biopharmaceutical industry location quotient	-0.025	-0.034	-0.016	-0.095
	(0.095)	(0.099)	(0.052)	(0.089)
Buyer- industry location quotient	1.607	-0.459	1.526	-0.646
	(12.692)	(1.208)	(4.338)	(0.535)
Supplier- industry location quotient	-0.058	0.028	-0.060	0.072
	(0.067)	(0.093)	(0.059)	(0.076)
Small firm birth rate	**0.142****	**0.181*****	**0.191*****	**0.183*****
	(0.069)	(0.064)	(0.055)	(0.064)
Number of small biopharmaceutical firms	0.103	0.069	**0.036****	0.021
	(0.110)	(0.061)	(0.017)	(0.013)
Square of number of small biopharmaceutical firms	-0.003	-0.002	**-0.000*****	**-0.000****
	(0.002)	(0.001)	(0.000)	(0.000)
Number of small biopharmaceutical firms per square mile	40.872	**48.172***	12.875	**31.303*****
	(37.733)	(25.291)	(24.902)	(10.648)
Time fixed effect		**1.054***		**0.856***
		(0.559)		(0.491)
Constant	-2.475	-1.566	**-3.044***	**-3.861*****
	(2.948)	(2.376)	(1.570)	(1.449)

Table 20 (continued)

	Catch-up regions		Leading and catch-up regions dataset	
Inflation variables	**Cross-Sectional Model**	**Panel Model**	**Cross-Sectional Model**	**Panel Model**
Life scientists job market share	-4.666 (3.279)	-1.935 (1.248)	**-4.955**** (1.971)	**-2.789***** (0.961)
Life scientists absolute salary ratio	-22.029 (34.415)	**-23.794**** (11.645)	**-18.548*** (11.268)	**-10.830***** (4.042)
Life scientists relative salary ratio	8.608 (8.280)	2.803 (11.717)	7.793 (6.994)	**7.076*** (3.790)
Number of biotech VC firms	**-12.367**** (5.017)	-3.084 (2.846)	**-18.594***** (1.907)	-1.092 (0.826)
Foreign born population share	0.329 (0.474)	0.047 (0.152)	0.266 (0.191)	0.044 (0.049)
Life science academic research share	-2.844 (7.891)	-7.792 (8.020)	-2.480 (2.527)	-1.742 (1.449)
Biopharmaceutical industry location quotient	-0.072 (0.443)	0.146 (0.492)	-0.079 (0.142)	-0.078 (0.152)
Buyer- industry location quotient	-3.325 (49.099)	0.503 (5.953)	-5.321 (23.218)	-1.317 (1.186)
Supplier- industry location quotient	-0.483 (0.355)	0.586 (0.719)	-0.495 (0.377)	0.168 (0.193)
Small firm birth rate	-0.343 (0.324)	0.164 (0.569)	-0.306 (0.198)	-0.067 (0.169)
Constant	15.750 (36.501)	**17.673**** (8.458)	13.492 (12.514)	4.565 (3.922)
Log likelihood	-160.483	-228.653	-208.131	-303.645
Overdispersion test (alpha =0)	9.58	10.05	127.88	134.03
Vuong test (ZINB vs.NBRM)	4.33	2.44	2.59	2.41
Number of nonzero observations	65	89	73	101
Number of zero observations	95	231	95	231

Note: Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

4.4. Summary

In this chapter, I find that the US biopharmaceutical industry has a relatively small size, plenty of high-paid jobs, high barrier for entry, and skewed spatial distribution. Over the last decade, there were a limited number of IPO and M&A events occurred in the biopharmaceutical industry, and most of them were concentrated in the leading regions.

The major findings from the regression analysis of the spatial distribution of biopharmaceutical IPO and M&A events are: first of all, there is strong evidence that local entrepreneurial climate is a key determinant of biopharmaceutical NTBF performance. This variable is consistently positive and significant in all model specifications. It is a key factor both for catch-up regions and leading regions; second, the impact of local life scientist job market size matters more in catch-up regions than in leading regions. In contrast, the salary ratio variables are more significant in leading regions than in catch-up regions; third, there is some evidence of the positive impact of proximity to VC firms on NTBF performance; fourth, there is little evidence of the significant impact of cultural diversity, academic research, industrial specialization, and coaggolomeration with buyer- or supplier- industries on NTBF performance; and finally, the adverse impact of proximity to too many biopharmaceutical firms is apparent in leading regions. .

CHAPTER 5: IT SERVICE INDUSTRY

This chapter analyzes the regional variations in the performance of Information Technology (IT) service NTBFs during the period from 1996 to 2005. The basic industry characteristics are discussed in Section 5.1. Section 5.2 describes the spatial distribution of NTBF IPO and acquisition activities in the industry. Section 5.3 uses a set of regression models to explore the determinant factors of the average NTBF performance in a region. Section 4 summarizes the main findings.

5.1. The IT Service Industry

In this study, the IT Service industry is defined by SIC 737 prior to 1997 and NAICS 5112 and 5415 thereafter. The IT service firms primarily engaged in Software Publishing, Computer Systems Design, Data Processing, and Other Computer Related Services[18].

5.1.1. Industry employment and earnings

The IT service industry is much larger than the biopharmaceutical industry in terms of employment (Table 21). The County Business Patterns data reveals that in 2005, the IT service industry had more than 1.5 million employees, accounting for 1.25 percent of total US employment. As reported in the previous chapter, the biopharmaceutical industry only had 247,847 employees in 2005, which is one sixth of the size of IT service industry.

From 1995 to 2005, the IT service industry employment increased by more than 30 percent, a gain of 342,004 new jobs. The industry reached its peak in 2001, when it employed more than 1.6 million employees, or with a share of national total employment of 1.4 percent.

[18] Source: Census Bureau, http://www.census.gov/epcd/naics02/def/NDEF325.HTM#N3254.

The burst of the Internet bubble in 2001 adversely affected the growth of the industry. Since 2002, the industry has grown no faster than the national average, and consequently its share of national employment has basically stagnated at 1.25 percent. In terms of the future growth, Hecker (2005) projected that the IT service industry will continue to grow and will add more than 50 percent new jobs from 2002-12 (p.59). This is in contrast to many high tech manufacturing industries in which employments are projected to decline.

Similar to the biopharmaceutical industry, the IT service industry provides high-paid jobs. The annual average salary in the IT service industry was at least twice the national annual average for all the years from 1995 to 2005. In 1999, it was even 3 times as high. Given its growth potential and high-paid jobs, the IT service industry is likely to continue to be a key source of economic growth for many regional economies.

Table 21: U.S. IT Service Industry Statistics: 1995-2005

Year	Employment	Share of total US employment (Percent)	Annual average salary	Ratio of industry to national annual average salary
1995	1,117,475	1.11	58,087	2.05
1996	1,266,890	1.24	62,707	2.16
1997	1,456,693	1.38	69,748	2.26
1998	1,156,452	1.07	81,197	2.54
1999	1,317,712	1.19	100,419	3.00
2000	1,502,721	1.32	95,012	2.69
2001	1,608,149	1.40	89,826	2.48
2002	1,401,599	1.25	86,497	2.35
2003	1,403,225	1.24	88,318	2.34
2004	1,433,721	1.25	86,529	2.20
2005	1,459,479	1.25	90,364	2.22

Source: (1) Data for employment were from the County Business Pattern, Bureau of the Census, Department of Commerce; (2) Data for average salary were from Quarterly Census of Employment and Salarys (QCEW), Bureau of Labor Statistics, Department of Labor.

5.1.2. Industry organization

The IT service industry is dominated by small firms. Table 22 reports that in 2005, firms with 1-99 employees accounted for 97.71 percent of total IT service establishments. The percentage for firms with 100-499 employees was 1.86. Firms with 500 employees or more only accounted for 0.23 percent. The size structure of the IT service industry is similar to the national pattern over the last decade. Compared to the biopharmaceutical industry, the IT service industry has much higher proportion of small firms and lower proportion of large companies. For example, as reported in Chapter 4, in 2005, the share for the size class 0-99, 100-499, and 500- in the biopharmaceutical industry was 75.93, 18.22, and 5.39 percent, respectively. Because the IT service has a lower barrier for new firms, which are typically small, the chance for the catch-up regions to catch up in the IT service industry should be higher than in the biopharmaceutical industry.

Table 22: Share of Total Establishments by Firm Size Classes for IT Service Industry and National Average: 1995-2005

Year	1~99		100~499		500~	
	IT Service	National	IT Service	National	IT Service	National
1995	97.22	97.72	2.51	2.03	0.27	0.25
1996	97.31	97.70	2.41	2.05	0.28	0.25
1997	97.39	97.68	2.34	2.07	0.26	0.25
1998	97.83	97.63	1.97	2.12	0.20	0.26
1999	97.68	97.57	2.11	2.17	0.22	0.26
2000	97.50	97.51	2.27	2.22	0.23	0.27
2001	97.33	97.51	2.39	2.21	0.28	0.27
2002	97.85	97.68	1.93	2.07	0.22	0.25
2003	97.88	97.67	1.90	2.09	0.23	0.25
2004	97.93	97.67	1.84	2.08	0.23	0.25
2005	97.91	97.66	1.86	2.09	0.23	0.25

Source: the County Business Pattern, Bureau of the Census, Department of Commerce.

5.1.3. Spatial distribution

The IT service industry is not as geographically concentrated as the U.S. biopharmaceutical industry. All the 51 states (including the District of Columbia) had IT service firms in 2004 (see Appendix 2), and 26 states had more than 10,000 employees in this industry. Only Wyoming had fewer than 1,000 employees. California had the largest share of 17.07 percent. It is followed by Virginia and Texas, with a share of 8.54 percent and 6.91 percent, respectively. The top ten states accounted for 62.89 percent of total IT service industry employment. As discussed in the previous chapter, the corresponding share for the biopharmaceutical industry was 71.81 percent, another indication that the biopharmaceutical industry is more concentrated than the IT service industry at the state level.

The spatial distribution of the IT service industry at state level has been relatively stable over time. Over the last decade, 27 states had higher shares in 2004 than in 1995, and 22 states had lower shares. However, the magnitude of increase or decrease is only moderate for most states. Washington experienced the largest share increase from 2.00 percent in 1995 to 4.07 percent in 2004. Illinois had the largest share drop, from 4.68 percent to 3.65 percent. Moreover, nine of the ten largest states in terms of IT service employment in 1995 continued to be ranked among top ten in 2004. Only Illinois fell out of the top ten, from the seventh in 1995 to the eleventh in 2004. The cumulative share for top ten states was 62.33 percent in 1995, which is only slightly lower than the 62.89 percent in 2004.

At the metropolitan area level, all 361 MSAs in 2004 had IT service establishments. Among them, 27 MSAs had more than 10,000 IT service employees, 87 MSAs had 1,000-10,000 employees, 47 MSAs had 500-1,000 employees, and the remaining 200 MSAs had fewer than 500 employees. Table 23 lists the twenty largest Metropolitan Areas in terms of industry employment in 2004. Washington-Arlington-Alexandria, New York-Northern New Jersey-Long Island, and San Francisco-Oakland-Fremont are the top three metropolitan areas, with a share of 10.18 percent, 7.0 percent, and 5.1 percent, respectively. The cumulative share for the top three metropolitan areas accounted for 22.28 percent of industry employment. This is compared to the 32.8 percent in the biopharmaceutical industry. Thus, the biopharmaceutical industry is more concentrated than the IT service industry even at the metropolitan area level.

Table 23: Twenty Largest IT Service Metropolitan Areas in 1995 and 2004

Metropolitan Statistical Areas	2004 Share	2004 Rank	2000 Share	2000 Rank	1995 Share	1995 Rank
Washington-Arlington-Alexandria, DC-VA-MD-WV	10.18	1	8.93	1	8.57	1
New York-Northern New Jersey-Long Island, NY-NJ-PA	7.00	2	8.46	2	8.03	2
San Francisco-Oakland-Fremont, CA	5.10	3	5.40	3	4.20	6
Boston-Cambridge-Quincy, MA-NH	4.76	4	5.09	4	5.33	3
San Jose-Sunnyvale-Santa Clara, CA	4.31	5	5.07	5	3.13	9
Los Angeles-Long Beach-Santa Ana, CA	4.31	6	4.03	7	5.22	4
Seattle-Tacoma-Bellevue, WA	3.65	7	3.05	11	2.56	11
Dallas-Fort Worth-Arlington, TX	3.34	8	3.25	10	3.98	7
Philadelphia-Camden-Wilmington, PA-NJ-DE-MD	3.23	9	3.30	9	2.93	10
Chicago-Naperville-Joliet, IL-IN-WI	3.21	10	4.48	6	4.27	5
Atlanta-Sandy Springs-Marietta, GA	2.83	11	3.40	8	3.17	8
Detroit-Warren-Livonia, MI	1.84	12	1.51	15	2.11	12
Minneapolis-St. Paul-Bloomington, MN-WI	1.78	13	1.98	12	2.10	13
Baltimore-Towson, MD	1.54	14	1.15	18	0.85	21
Houston-Baytown-Sugar Land, TX	1.51	15	1.59	14	1.63	15
San Diego-Carlsbad-San Marcos, CA	1.48	16	1.38	16	1.27	16
Miami-Fort Lauderdale-Miami Beach, FL	1.32	17	1.07	20	0.65	33
Denver-Aurora, CO	1.30	18	1.79	13	1.70	14
Phoenix-Mesa-Scottsdale, AZ	1.18	19	1.11	19	0.78	25
Austin-Round Rock, TX	1.18	20	1.22	17	0.66	31

Source: the County Business Pattern, Bureau of the Census, Department of Commerce.

In terms of the changes in industry share from 1995 to 2004, 109 of 361 MSAs had higher shares in 2004 than in 1995, 134 had lower shares, and the remaining 118 had the same share of industry employment[19]. Seattle-Tacoma-Bellevue (WA) had experienced the largest

[19] One should be cautious when interpreting the industrial employment changes at the metropolitan level because the CBP data is not completely comparable over time.

share increase, followed by Washington-Arlington-Alexandria (DC-VA-MD-WV) and San Jose-Sunnyvale-Santa Clara (CA). All of these three metropolitan areas saw increases of more than one share point. In contrast, Chicago-Naperville-Joliet (IL-IN-WI), New York-Northern New Jersey-Long Island (NY-NJ-PA), and Los Angeles-Long Beach-Santa Ana (CA) had the largest share drops, ranging from -1.06 to -1.90 share point.

Because this study defines a MSA as a 'leading' region if its high tech index value in 1995 was ranked 95th percentile or above, 15 MSAs are classified as leading regions, and the other 301MSAs are categorized as catch-up regions in the IT service industry. Table 24 reports that in 1995, the 301 catch-up MSAs accounted for 42.68 percent of total IT service employment, 48.12 percent of total IT service establishments, and 45.06 percent of small and medium IT service establishments. From 1995 to 2000, the catch-up regions had lost some share in all these three measures.

Table 24: Share of IT Service Industry for Leading and Catch-Up Regions, in 1995 and 2000

	1995			2000		
	Emp. Share	Estab. Share	SME Estab. Share	Emp. Share	Estab. Share	SME Estab. Share
Leading regions	57.32	51.88	54.94	59.88	57.86	57.84
Catch-up regions	42.68	48.12	45.06	40.12	42.14	42.16

Source: the County Business Pattern, Bureau of the Census, Department of Commerce.

5.2. IPO and M&A Activities in the IT Service Industry

The IPO and M&A activities were much more active in the IT service industry than in the biopharmaceutical industry during the period from 1996 to 2005. Over this period, 636 IT service firms launched their Initial public offerings. Most qualify as NTBFs: 414 had fewer than 500 employees and were younger than 20 years when they became public. Figure 7 shows the number of total and small IT service IPOs each year. It suggests that the IPO market for the IT Service industry also fluctuated over time. There were far more IT service IPOs in the first five years than in the second five years. This is because after the burst of Internet bubble in 2000, the IT service firms became extremely unpopular in the equity market.

In terms of M&A activities, there were 7868 deals completed during the study period. Because 1170 of them had no location information, this analysis is restricted to the remaining 6698 deals. Figure 8 shows the number of missing cases is roughly proportional to the total number of acquisition activities each year. Therefore, it may not cause serious bias over time.

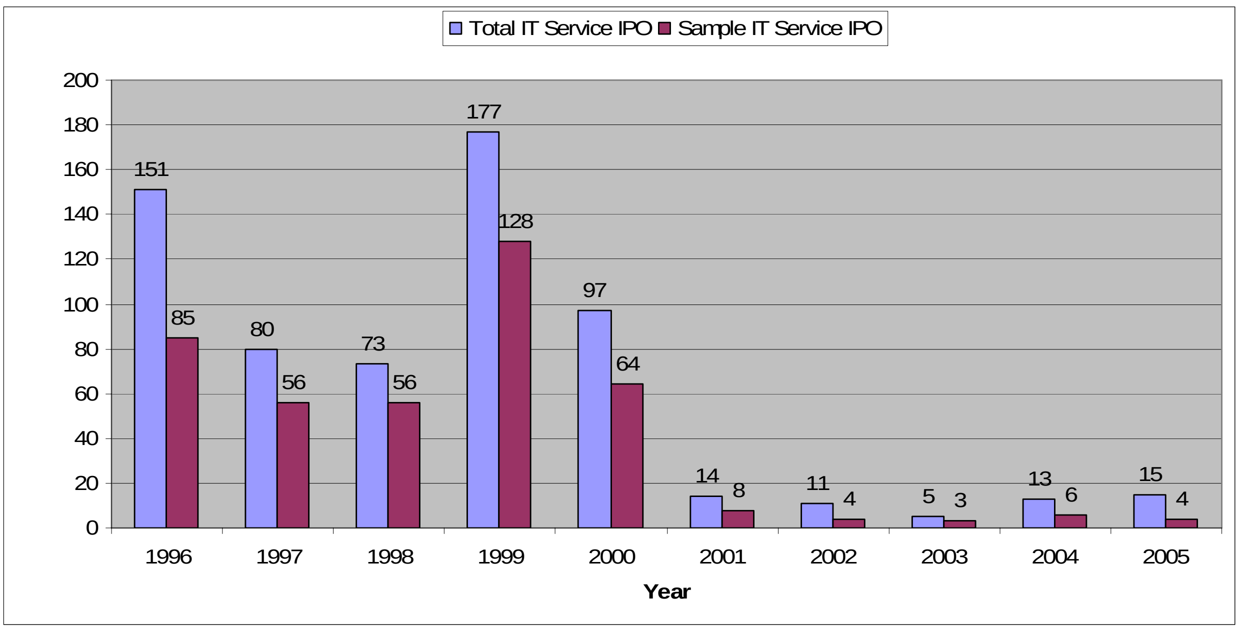


Figure 7: Number of Total and Sample IT Service IPO Events, 1996- 2005

Source: Securities Data Company, New Issues Database; Hoover's IPO list; Jay Ritter's 1975-2005 IPO dataset; and University of Chicago's Center for Research in Security Prices database.

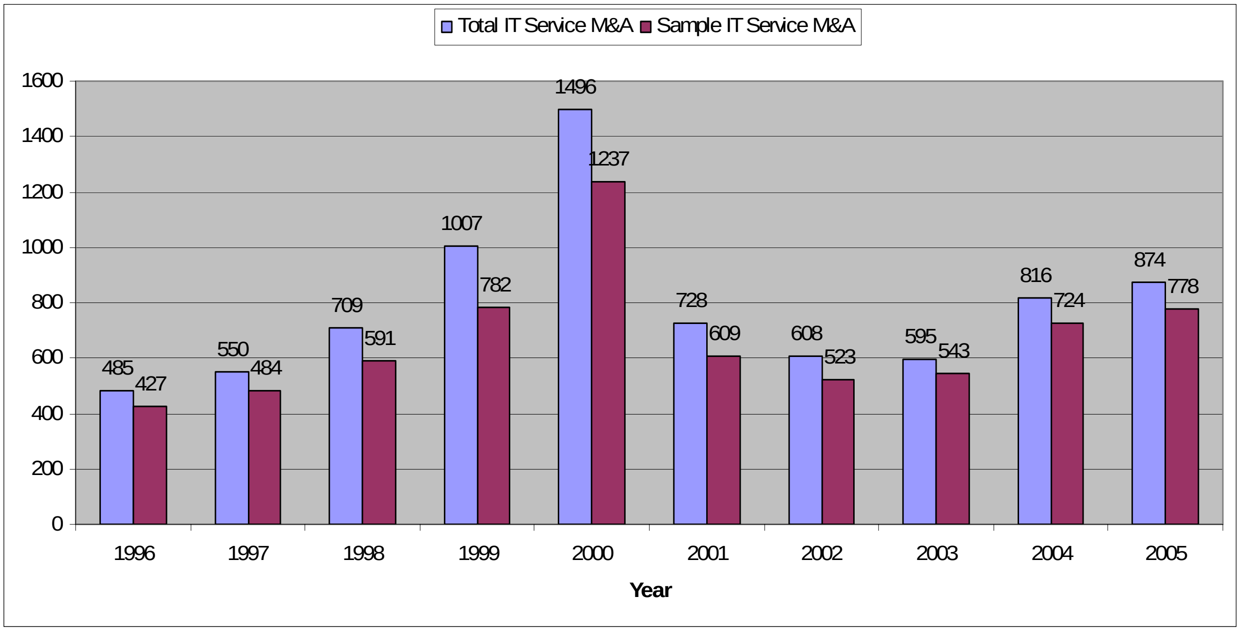


Figure 8: Number of Total and Sample IT Service M&A Events, 1996- 2005

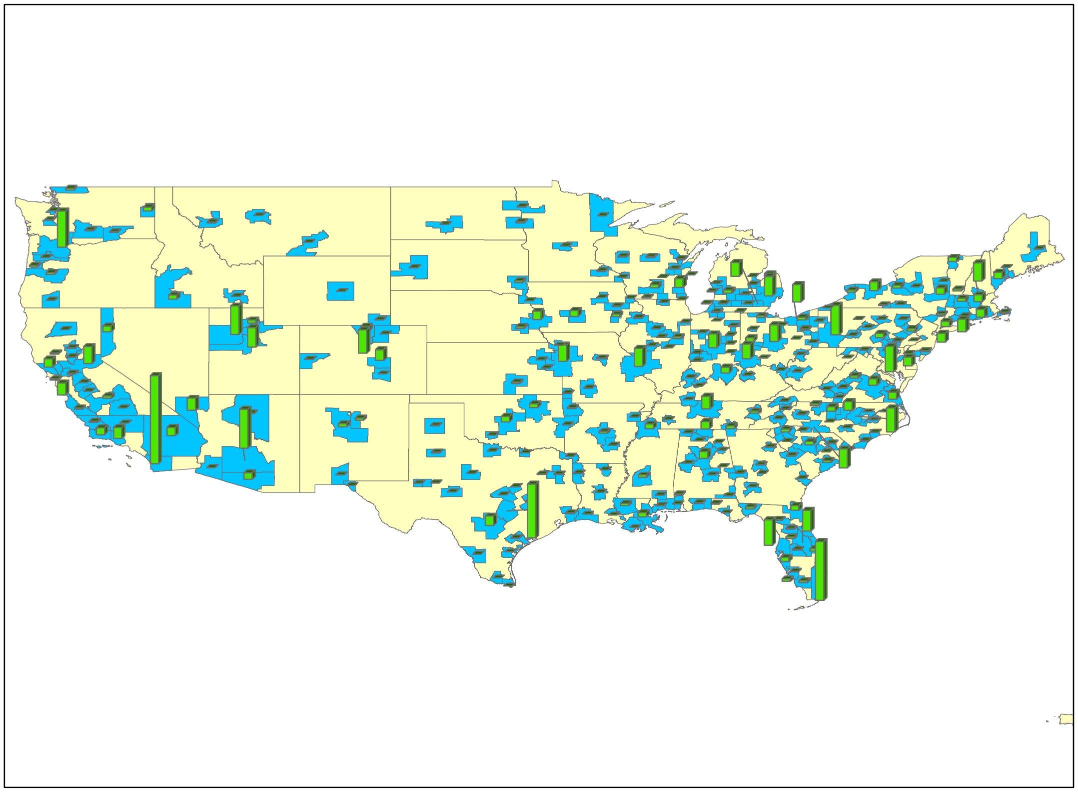
Source: Securities Data Company, New Issues Database.

The spatial distribution of the IT service IPO and M&A events is highly skewed. Table 25 reports that 4583, or 65.64 percent of total events occurred in the 15 leading MSAs. The 301 catch-up MSAs only had 2399, or 34.35 percent, much lower than its share of industrial employment, which was 42.68 percent in 1995. The top five regions in terms of number of IT service IPO and M&A events are: San Francisco-Oakland-Fremont (CA), New York-Northern New Jersey-Long Island (NY-NJ-PA), San Jose-Sunnyvale-Santa Clara (CA), Boston-Cambridge-Quincy (MA-NH), and Los Angeles-Long Beach-Santa Ana (CA). These five regions in total had 2756, or 39.47 percent of total deals. Among the 301 catch-up regions, 199 MSAs, or 66.12 percent, had one or more deals. San Diego-Carlsbad-San Marcos (CA) was the leading catch-up region, with 189 deals. Miami-Fort Lauderdale-Miami Beach (FL), Austin-Round Rock (TX), and Phoenix-Mesa-Scottsdale (AZ) followed, with 126, 115, and 84, respectively. Figure 9 depicts the spatial distribution of IT service IPO and M&A events in catch-up Regions.

Table 25: Statistics of IT Service IPO and M&A Events for the Leading and Catch-Up Regions

	Sum	Mean	Max	Min	Std
Leading regions	4583	305.53	697	53	197.88
Catch-up regions	2399	7.97	189	0	19.07

Figure 9: Spatial Distribution of IT Service IPO and M&A Events in Catch-Up Regions



5.3. Determinants of Regional Variations in the IT Service NTBF Performance

5.3.1. Models

This study employs both the cross-sectional and two-period panel ZINB models to identify the causal factors that affect the average IT service NTBF performance in a catch-up region. The cross-section models predict the number of IT service IPO and acquisition events that occurred in a region between 1996 and 2005 by a set of regional characteristics in 1995 or other years. These models are limited to the 301 catch-up MSAs where the IT service industry had consistently existed from 1995 to 2004 so that all the observations had the same exposure time to the events. The two-period panel data models predict the number of IT service IPO and M&A events occurred in a region during two sub periods (1996-2000 and 2001-2005). The values for the predictor variables were updated every five years. The sample size for the panel data model is 602.

The Breusch-Pagan / Cook-Weisberg heteroskedasticity test statistic is 2362.90 ($p<0.00$) for the catch-up cross-sectional model, and 4681.21 ($p<0.00$) for the catch-up two-period panel model, suggesting the presence of heteroskedasticity in the data. Therefore, this study only reports the robust estimation results.

Table 26 and 27 report the descriptive statistics and table 28 and 29 presents the Pearson correlation coefficient matrix both for the cross-sectional and panel data. As in the biopharmaceutical industry, some correlation coefficients are very large, indicating that multicollinearity might be a concern in the data. For example, Table 28 reports that in the cross-sectional dataset, the control variable, a region's total job market size (emp_share), is strongly correlated with its computer scientist job market size (job_share), with a correlation coefficient of 0.95. In addition, the correlation coefficient between the absolute salary level of life scientists (Salary_log) and relative salary level (Relative_salary) is 0.72. The presence of multicollinearity doesn't affect the overall predictive power or the statistical significance of the model. However, it can increase the standard errors of the regression coefficients and consequently underestimate their significant level.

As in the biopharmaceutical industry, this study employs two approaches to address the potential multicollinearity problem. First, the control variable, a region's total job market size, is excluded from the full model. Second, besides the full model, three reduced models defined in Chapter 4 are adopted to distinguish the total effect of two computer scientists salary level variables on firm performance from their direct effect (when hold the other variables constant). Specifically, Reduced model 1 only includes the absolute salary level variable (Absolute_salary). Reduced model 2 only includes the relative salary level variable (Relative_salary). Reduced model 3 also only includes the relative salary variable, but it is based upon a smaller dataset in which 15 outlier regions were excluded. Regions with high absolute salary level are typically with a high relative salary level because they are highly correlated. However, in some cases, regions with low absolute salary level can also have high relative salary level because the denominator, the average salary level for all occupational jobs are very low. The presence of such outlier regions may completely distort the causal relationship between the life scientist relatively salary ratio and firm performance. To identify these outlier regions, I calculate the percentile ranks based upon the absolute and relative salary level of local life scientists, respectively. Seven regions were defined as the outliers because the difference in the percentile ranks for these regions is over 60. See appendix 3 for the results of these three reduced models.

All the model specifications use the same set of 'inflate' variables that predict the probability of being in the 'always-0-group' for a region. For the panel data, all the independent and control variables except their squared terms are used as the predictors of the odds of 'being-always-0-group.

Table 26: Descriptive Statistics of IT Service Cross-Sectional Variables (N=301)

Variable	Mean	Std Dev	Minimum	Maximum
IPO_Target	7.97	19.07	0.00	189.00
Job_share	0.15	0.39	0.00	5.04
Absolute_salary	0.99	0.13	0.62	1.35
Relative_Salary	1.53	0.17	0.91	2.19
VC_firm	0.48	1.47	0.00	12.00
VC_firm _sq	2.37	12.80	0.00	144.00
Immi_share	4.22	4.56	0.39	25.00
Univ_share	0.18	0.86	0.00	10.95
Ind_lq	0.37	0.40	0.01	2.88
Buyer_lq	0.59	2.93	0.00	33.21
Supplier_lq	2.23	4.02	0.00	53.97
Firm_Birth_Rate	1.17	2.10	6.60	18.80
Ind_SME	105.94	176.40	1.00	1097.00
Ind_SME _sq	42237.56	145474.03	1.00	1203564.00
Est_density	0.05	0.09	0.00	1.02
Emp_share	0.25	0.33	0.04	2.38

Table 27: Descriptive statistics for variables in the IT service panel data models (N=602)

Variable	Mean	Std Dev	Minimum	Maximum
IPO_Target	3.99	9.69	0.00	108.00
Job_share	0.17	0.41	0.00	6.26
Absolute_salary	0.99	0.13	0.62	1.35
Relative_Salary	1.63	0.20	0.91	2.56
VC_firm	0.43	1.32	0.00	12.00
VC_firm _sq	1.93	11.22	0.00	144.00
Immi_share	5.15	5.25	0.39	31.21
Univ_share	0.19	0.85	0.00	10.95
Ind_lq	0.35	0.40	0.01	3.71
Buyer_lq	0.55	3.02	0.00	40.91
Supplier_lq	2.04	3.32	0.00	53.97
Firm_Birth_Rate	1.13	2.10	6.20	21.30
Ind_SME	126.39	226.71	1.00	2192.00
Ind_SME _sq	67285.43	287502.06	1.00	4804864.00
Est_density	0.06	0.12	0.00	1.72
Emp_share	0.25	0.33	0.04	2.43
Period_dummy	3.99	9.69	0.00	108.00

Table 28: Correlation Coefficient Matrix for the IT Service Cross-Sectional Data

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
(1) IPO_Target	1.00													
(2) Job_share	0.48	1.00												
	<.0001													
(3) Absolute_salary	0.36	0.34	1.00											
	<.0001	<.0001												
(4) Relative_salary	0.14	0.14	**0.72**	1.00										
	0.01	0.02	<.0001											
(5) VC_firm	**0.70**	0.43	0.29	0.12	1.00									
	<.0001	<.0001	<.0001	0.04										
(6) Immi_share	0.27	0.08	0.19	0.01	0.07	1.00								
	<.0001	0.17	0.00	0.92	0.22									
(7) Univ_share	0.36	0.21	0.17	0.05	0.64	0.03	1.00							
	<.0001	0.00	0.00	0.41	<.0001	0.57								
(8) Ind_lq	0.41	0.24	0.32	0.10	0.32	0.07	0.15	1.00						
	<.0001	<.0001	<.0001	0.10	<.0001	0.22	0.01							
(9) Buyer_lq	0.11	0.14	0.14	0.11	0.33	-0.01	0.08	0.25	1.00					
	0.05	0.01	0.02	0.06	<.0001	0.80	0.17	<.0001						
(10) Supplier_lq	0.12	0.07	0.07	0.02	0.08	0.01	0.03	0.46	0.06	1.00				
	0.03	0.21	0.20	0.71	0.16	0.80	0.57	<.0001	0.31					
(11)Firm_Birth_Rate	0.32	0.12	0.15	0.22	0.16	0.27	0.00	0.16	0.14	0.10	1.00			
	<.0001	0.03	0.01	<.0001	0.01	<.0001	1.00	0.01	0.02	0.10				
(12) Ind_ SME	**0.85**	0.56	0.41	0.14	0.63	0.20	0.39	0.41	0.06	0.18	0.22	1.00		
	<.0001	<.0001	<.0001	0.02	<.0001	0.00	<.0001	<.0001	0.28	0.00	0.00			
(13) Est_density	0.45	0.29	0.40	0.06	0.36	0.19	0.25	0.42	0.14	0.20	0.05	0.54	1.00	
	<.0001	<.0001	<.0001	0.30	<.0001	0.00	<.0001	<.0001	0.01	0.00	0.41	<.0001		
(14) Emp_share	**0.75**	0.53	0.36	0.11	0.56	0.22	0.35	0.30	-0.01	0.14	0.17	**0.95**	0.44	1.00
	<.0001	<.0001	<.0001	0.06	<.0001	0.00	<.0001	<.0001	0.86	0.02	0.00	<.0001	<.0001	

Note: a. The square terms are not included in order to save space; b. the first contains the correlation coefficient, and the p-values are in the second row.
b: number of observations: 301

Table 29: Correlation Coefficient Matrix for the IT Service Panel Data

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
(1) IPO_Target	1.00													
(2) Job_share	0.46	1.00												
	<.0001													
(3) Absolute_salary	0.37	0.32	1.00											
	<.0001	<.0001												
(4) Relative_salary	0.06	0.05	0.54	1.00										
	0.12	0.20	<.0001											
(5) VC_firm	0.66	0.40	0.31	0.03	1.00									
	<.0001	<.0001	<.0001	0.42										
(6) Immi_share	0.28	0.13	0.30	0.22	0.08	1.00								
	<.0001	0.00	<.0001	<.0001	0.04									
(7) Univ_share	0.42	0.23	0.19	0.03	0.64	0.07	1.00							
	<.0001	<.0001	<.0001	0.50	<.0001	0.11								
(8) Ind_lq	0.44	0.25	0.37	-0.01	0.38	0.09	0.21	1.00						
	<.0001	<.0001	<.0001	0.84	<.0001	0.03	<.0001							
(9) Buyer_lq	0.09	0.10	0.11	0.05	0.26	0.01	0.07	0.23	1.00					
	0.03	0.02	0.01	0.20	<.0001	0.75	0.11	<.0001						
(10) Supplier_lq	0.12	0.10	0.09	-0.04	0.11	-0.02	0.03	0.40	0.03	1.00				
	0.00	0.01	0.02	0.31	0.01	0.70	0.40	<.0001	0.43					
(11)Firm_Birth_Rate	0.33	0.12	0.19	0.08	0.16	0.35	0.02	0.22	0.08	0.08	1.00			
	<.0001	0.00	<.0001	0.04	<.0001	<.0001	0.64	<.0001	0.04	0.04				
(12) Ind_ SME	0.82	0.53	0.40	0.10	0.59	0.27	0.39	0.42	0.04	0.16	0.23	1.00		
	<.0001	<.0001	<.0001	0.02	<.0001	<.0001	<.0001	<.0001	0.30	<.0001	<.0001			
(13) Est_density	0.43	0.27	0.40	0.03	0.35	0.22	0.24	0.46	0.08	0.16	0.08	0.54	1.00	
	<.0001	<.0001	<.0001	0.47	<.0001	<.0001	<.0001	<.0001	0.04	0.00	0.06	<.0001		
(14) Emp_share	**0.75**	0.51	0.36	0.05	0.55	0.24	0.34	0.31	-0.01	0.15	0.20	**0.92**	0.41	1.00
	<.0001	<.0001	<.0001	0.23	<.0001	<.0001	<.0001	<.0001	0.78	0.00	<.0001	<.0001	<.0001	

Note: a. The square terms are not included in order to save space; b. the first contains the correlation coefficient, and the p-values are in the second row.
b: number of observations: 602

5.3.2 Results

This section presents the results for the empirical estimation of the effects of regional characteristics on the average IT service NTBF performance in a region, measured by the number of IPO and M&A events. I first report the estimation results from the cross-sectional models. The temporal stability test findings based upon two-period panel data models are followed. Finally, I present the results from the distance-weighted dataset, which is created to capture the 'neighboring' effect.

(1) Cross-sectional models

Table 30 presents the coefficient estimates of the cross-sectional zero-inflated negative models both for catch-up regions and all the regions. In each model, the top set of coefficients corresponds to the NBRM for those in the 'Not- Always-0-Group'. The lower set of coefficients, labeled 'inflate', corresponds to the binary logistic model predicting membership in the group that always has zero counts. Table 30 also reports the results of alpha and Vuong test in the end. The alpha test in each model is larger than 1.96, which provides strong evidence of overdispersion. Therefore, the negative binomial regression model is preferred to the Poisson regression model in this data. The Vuong test compares the zero-inflated regression model specification with the standard negative binomial regression model. The Vuong test statistic in each model is larger than 1.96, indicating that the zero-inflated regression model specification is favored in this data.

The results provide less evidence of the significance of local scientist job market size in the IT service industry than in the biopharmaceutical industry. The estimated coefficient of this variable is negative and insignificant for catch-up regions. Multicolinearity is unlikely to be the cause for the lack of the significance of this variable because it has only modest correlation coefficient with other variables. The correlation coefficient matrix (Table 28) shows that this variable has the strongest correlation with number of small IT service firms (Ind_sme_lg), but the correlation coefficient is only 0.56. When the 15 leading regions are included into the analysis, the variable became negative and significant ($p<0.05$). One explanation for the negative impact of local computer scientist job market is that the spatial distribution of computer scientist jobs is less concentrated than that of life scientist jobs. In terms of predicting the odds of being the 'always-0-group', this variable is negative and insignificant.

The absolute salary level of computer scientists in a region is found to have positive and significant ($p<0.05$) impact on NTBF performance, but the effect of relative salary level is negative and insignificant. The empirical evidence again suggests that it is the absolute salary rate rather than the relative salary rate that can positively affect the performance of local NTBFs in catch-up regions. It indicates that regions that can provide nationally competitive high-paid jobs for computer scientists are more likely to be associated with better NTBF performance. The negative impact of relative salary ratio is significant ($p<0.001$) only in the full sample data. In terms of predicting the odds of being the 'always-0-group', the absolute salary ratio is negative and insignificant, and the relative salary ratio is positive and insignificant.

The results report that the nonlinear effect of proximity to VC firms is more apparent in the IT service industry than in the biopharmaceutical industry. For catch-up regions, the number of IT service VC firms has a positive and highly significant ($p<0.01$) coefficient. Its squared term is also significant ($p<0.05$), but with a negative sign. The model predicts that the negative impact began to dominate the positive impact of proximity to specialized VC firms when there are 7 or more VC firms in a metropolitan area. For the full sample model in which both the leading and catch-up regions are included, the coefficient of number of IT service VC firm even became negative and significant ($p< 0.05$). This finding is consistent with previous studies suggesting the adverse effects of 'excess' venture capital. Some literature has argued that the excess amount of venture capital in a region is likely to flow straight to low-quality entrepreneurship(Acs and Storey 2004; Venkataraman 2004). Stuart and Sorenson (2003) pointed out that proximity to many VCs implies that a focal organization may compete against well-financed rivals in local factor market.

Cultural diversity is found to have positive and significant ($p<0.05$) impact on the performance of IT service NTBF performance in catch-up regions. This finding supports Florida's (2002) argument that place can attract talented people by cultural diversity or openness to all lines of lifestyles. However, an alternative explanation is that the IT service industry hires more foreign-born scientists and engineers than the biopharmaceutical industry does.

In contrast to the biopharmaceutical industry, the industrial specialization has significant ($p<0.01$) and positive effect on NTBF performance. This finding provides evidence of positive spillovers associated with proximity to similar firms. The impact of buyer industries is negative and insignificant. Coaggolomeration with supplier-industries has negative and significant ($p<0.1$) effect. Therefore, the results are not completely consistent with Porter's cluster theory (1990, 1998).

The results provide little evidence of the direct and significant impact of academic computer science research on the performance of local small IT service firms in catch-up regions. However, when the 15 IT service leading regions are included into the analysis, this variable became positive and significant ($p<0.05$). This finding suggests that academic research plays a more important role in the leading regions than in catch-up regions.

As in the biopharmaceutical industry, the findings provide strong evidence that the overall entrepreneurial climate in a region has significant and positive impacts on the performance of IT service NTBFs. The local entrepreneurial climate, measured by the small firm birth rate, is positive and highly significant ($p<0.01$) both in catch-up regions and the full sample data.

The control variables, number of small IT service firms its squared term, are significant both in the catch-up and full sample model. The number of small IT service firms is positive and its squared term is negative. Both variables are significant at 0.05 level in catch-up regions and 0.01 level in the full sample model. The model predicts when the size of impact of industrial clusters on firm performance reaches the maximum when there are about 742 small IT service in firms in catch-up regions.

Table 30: Results of Cross-Sectional ZINB Models of IT Service IPO and M&A Events between 1996 and 2005

Independent variables	Catch-up regions (N=301)	Leading and catch-up regions (N=316)
Computer scientists job market share	-0.136 (0.140)	**-0.418**** (0.199)
Computer scientists absolute salary ratio	**1.952**** (0.961)	**5.142***** (1.018)
Computer scientists relative salary ratio	-0.734 (1.090)	**-1.932***** (0.724)
Number of IT service VC firms	**0.137***** (0.049)	**-0.041**** (0.019)
Square of number of IT service VC firms	**-0.010**** (0.004)	0.000 (0.000)
Foreign born population share	**0.036**** (0.018)	-0.012 (0.017)
Computer science academic research share	0.031 (0.027)	**0.130**** (0.064)
IT service industry location quotient	**0.476***** (0.093)	**0.677***** (0.154)
Buyer- industry location quotient	-0.005 (0.034)	0.010 (0.015)
Supplier- industry location quotient	**-0.011*** (0.007)	-0.005 (0.011)
Small firm birth rate	**0.143***** (0.025)	**0.189***** (0.029)
Number of small IT service firms	**0.008**** (0.003)	**0.003***** (0.000)
Square of number of small IT service firms	**-0.000**** (0.000)	**-0.000***** (0.000)
Number of small IT service firms per square mile	-0.152 (0.351)	0.247 (0.375)
Constant	**-2.289***** (0.532)	**-3.552***** (0.744)

Table 30 (continued)

Inflation variables	**Catch-up regions (N=301)**	**Leading and catch-up regions (N=316)**
Computer scientists job market share	-0.596 (0.534)	-1.183 (2.541)
Computer scientists absolute salary	-7.435 (8.873)	-4.543 (4.359)
Computer scientists relative salary	5.361 (3.591)	3.592 (2.776)
Number of IT service VC firms	0.535 (3.204)	1.191 (1.146)
Foreign born population share	0.103 (0.110)	0.029 (0.064)
Computer science academic research share	1.611 (3.795)	**3.432*** (1.888)
IT service industry location quotient	-0.189 (1.088)	-0.120 (1.168)
Buyer- industry location quotient	0.026 (0.281)	0.105 (0.094)
Supplier- industry location quotient	-0.131 (0.141)	-0.141 (0.124)
Small firm birth rate	-0.126 (0.472)	0.033 (0.160)
Number of small IT service firms	-0.082 (0.240)	**-0.172***** (0.048)
Number of small IT service firms per square mile	16.417 (54.581)	**33.790***** (10.451)
Constant	1.056 (2.794)	0.484 (3.336)
Log likelihood	-576.923	-652.489
Overdispersion test (alpha =0)	130.53	1716.31
Vuong test (ZINB vs. NBRM)	3.73	3.46
Number of nonzero observations	199	214
Number of zero observations	102	102

Note: Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

(2) Two-period panel data models

Table 32 reports the coefficient estimates of the two-period zero-inflated negative models both for catch-up regions and the full sample which consists of both the leading and catch-up regions. Again, the overdispersion and Vuong test statistics suggest that the ZINB specification is the best choice for this data.

The results from the two-period panel data models indicate the following causal patterns are persistent over time: (1) the impact of the absolute salary level of computer scientists is significant and positive, but the effect of relative salary level is insignificant and negative; (2) proximity to IT service venture capital firms has significantly positive but nonlinear effect; (3) industrial specialization has positive and significant impacts; (4) the overall entrepreneurial climate in a region has positive and highly significant impact on IT service NTBFs performance; and (5) the number of small IT service firms has significantly positive and non-linear impact.

Compared to the cross-sectional design, the two-period panel models provide less evidence of the significance of cultural diversity. The estimated coefficient of foreign born population is positive but insignificant both in the catch-up and full sample model.

Table 31: Results of Two-Period Panel ZINB Models of IT Service IPO and M&A Events between 1996 and 2005

Independent variables	Catch-up regions (N=602)	Leading and catch-up regions (N=632)
Computer scientists job market share	-0.101 (0.108)	-0.040 (0.184)
Computer scientists absolute salary ratio	**1.993*** (1.071)	**4.774***** (0.762)
Computer scientists relative salary ratio	-0.098 (0.724)	**-1.236**** (0.622)
Number of IT service VC firms	**0.224***** (0.052)	0.004 (0.020)
Square of number of IT service VC firms	**-0.017***** (0.005)	-0.000 (0.000)
Foreign born population share	0.014 (0.013)	-0.015 (0.015)
Computer science academic research share	0.010 (0.031)	**0.090**** (0.046)
IT service industry location quotient	**0.450***** (0.098)	**0.464***** (0.154)
Buyer- industry location quotient	-0.020 (0.014)	**-0.021*** (0.012)
Supplier- industry location quotient	-0.011 (0.010)	0.006 (0.020)
Small firm birth rate	**0.101***** (0.019)	**0.155***** (0.025)
Number of small IT service firms	**0.004***** (0.001)	**0.001***** (0.000)
Square of number of small IT service firms	**-0.000***** (0.000)	**-0.000***** (0.000)
Number of small IT service firms per square mile	-0.102 (0.236)	0.118 (0.316)
Time fixed effect	-0.168 (0.247)	0.183 (0.146)
Constant	**-2.777***** (0.696)	**-3.950***** (0.778)

Table 31 (continued)

Inflation variables	Catch-up regions (N=602)	Leading and catch-up regions (N=632)
Computer scientists job market share	-0.061 (0.261)	-0.025 (0.251)
Computer scientists absolute salary	-2.932 (2.182)	-1.715 (2.311)
Computer scientists relative salary	**3.669***** (1.392)	**3.425**** (1.349)
Number of IT service VC firms	0.243 (0.532)	-0.011 (0.688)
Foreign born population share	0.015 (0.032)	-0.014 (0.040)
Computer science academic research share	-0.060 (0.243)	-0.025 (0.233)
IT service industry location quotient	**1.510*** (0.825)	**1.421**** (0.692)
Buyer- industry location quotient	**-0.273*** (0.143)	**-0.252*** (0.137)
Supplier- industry location quotient	-0.176 (0.108)	-0.152 (0.108)
Small firm birth rate	**-0.175*** (0.095)	**-0.163*** (0.096)
Number of small IT service firms	**-0.049***** (0.017)	**-0.051***** (0.012)
Number of small IT service firms per square mile	3.907 (7.818)	4.734 (7.455)
Constant	0.225 (2.195)	-0.544 (2.233)
Log likelihood	-863.451	-1170.906
Overdispersion test (alpha =0)	384.44	2587.92
Vuong test (ZINB vs. NBRM)	2.35	5.02
Number of nonzero observations	319	349
Number of zero observations	283	283

Note: Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

(3) Distance-weighted results

This study replicates the cross-sectional and two-period models to a set of distance-weighted variables. The following variables are recalculated using the method defined in Chapter 3: computer scientist job market share, number of IT service VC firms and its squared term, foreign born population share, computer science academic research, number of small IT service firms, and its squared term. These variables are likely to have impacts on the neighboring metropolitan areas. Table 33 presents the results. Again, the overdispersion and Vuong test statistics in each model are larger than 1.96, indicating that the ZINB specification is favored.

The distance-weighted models provide further evidence supporting the following causal relationships: (1) the job market size of computer scientist has negative and significant impact in both the catch-up and full sample models; (2) the impact of the absolute salary level of computer scientist is positive and highly significant ($p<0.01$), but the effect of relative salary level is negative and significant ($p<0.1$ in the catch-up cross-sectional model and $p<0.05$ or 0.01 in the full sample models; (3) proximity to the IT service venture capital firms has positive but non-linear impact; (4) cultural diversity is positive and highly significant in catch-up region models; (5) the positive impact of academic life science research on NTBF performance is significant only when the leading regions are included into the analysis; (6) industrial specialization is found to have positive and highly significant impact; (7) the overall entrepreneurial climate in a region has positive and highly significant impact; (8) the number of small IT service firms has positive and non-linear impact. Such a non-linear impact is apparent both in IT service catch-up and leading regions.

Table 32: Distance-Weighted Results of ZINB Models of IT Service IPO and M&A Events between 1996 and 2005

	Catch-up regions		Leading and catch-up regions	
Inflation variables	**Cross-Sectional (N=301)**	**Panel (N=602)**	**Cross-Sectional (N=316)**	**Panel (N=632)**
Computer scientists job market share	**-0.173**** (0.083)	-0.106 (0.073)	**-0.272*** (0.140)	-0.054 (0.103)
Computer scientists absolute salary ratio	**2.083***** (0.618)	**2.120***** (0.787)	**5.129***** (1.039)	**5.016***** (0.808)
Computer scientists relative salary ratio	**-0.869*** (0.486)	-0.199 (0.618)	**-1.939***** (0.732)	**-1.382**** (0.648)
Number of IT service VC firms	**0.056*** (0.031)	**0.136***** (0.035)	**-0.027*** (0.014)	0.005 (0.014)
Square of number of IT service VC firm	-0.003 (0.003)	**-0.010***** (0.003)	0.000 (0.000)	-0.000 (0.000)
Foreign born population share	**0.038***** (0.010)	**0.021**** (0.009)	-0.000 (0.013)	-0.006 (0.011)
Computer Science academic research share	0.001 (0.021)	-0.001 (0.018)	**0.075**** (0.037)	**0.055**** (0.027)
IT service industry location quotient	**0.496***** (0.089)	**0.454***** (0.101)	**0.696***** (0.163)	**0.480***** (0.164)
Buyer- industry location quotient	-0.004 (0.017)	**-0.022*** (0.013)	0.011 (0.016)	**-0.022*** (0.012)
Supplier- industry location quotient	-0.010 (0.006)	-0.008 (0.011)	-0.005 (0.011)	0.008 (0.021)
Small firm birth rate	**0.179***** (0.038)	**0.133***** (0.017)	**0.202***** (0.029)	**0.173***** (0.271)
Number of small IT service firms	**0.006***** (0.001)	**0.003***** (0.000)	**0.002***** (0.000)	**0.001***** (0.000)
Square of number of small IT service firms	**-0.000***** (0.000)	**-0.000***** (0.000)	**-0.000***** (0.000)	**-0.000***** (0.000)
Number of small IT service firms per square mile	-0.573 (0.378)	-0.346 (0.216)	0.047 (0.204)	0.037 (0.193)
Time fixed effect		-0.178 (0.214)		0.208 (0.150)
Constant	**-2.735***** (0.612)	**-3.224***** (0.616)	**-3.746***** (0.774)	**-4.225***** (0.817)

Table 32 (continued)

Inflation variables	Catch-up regions Cross-Sectional (N=301)	Catch-up regions Panel (N=602)	Leading and catch-up regions Cross-Sectional (N=316)	Leading and catch-up regions Panel (N=632)
Computer scientists job market share	-0.355 (0.367)	-0.050 (0.157)	-0.625 (1.326)	-0.009 (0.151)
Computer scientists absolute salary ratio	-7.176 (6.002)	-2.625 (2.194)	-4.111 (4.430)	-1.419 (2.321)
Computer scientists relative salary ratio	5.185 (3.877)	**3.417**** (1.334)	3.627 (2.836)	**3.246**** (1.329)
Number of IT service VC firms	0.463 (0.922)	0.177 (0.380)	0.746 (0.996)	-0.014 (0.491)
Foreign born population share	0.072 (0.056)	0.012 (0.023)	0.015 (0.043)	-0.011 (0.027)
Computer science academic research share	1.459 (1.584)	0.011 (0.152)	**2.505*** (1.286)	0.025 (0.139)
IT service industry location quotient	-0.429 (1.170)	**1.267*** (0.755)	-0.411 (1.235)	**1.169*** (0.671)
Buyer- industry location quotient	0.048 (0.133)	**-0.262*** (0.140)	0.102 (0.093)	**-0.254*** (0.140)
Supplier- industry location quotient	-0.104 (0.108)	-0.154 (0.100)	-0.123 (0.125)	-0.136 (0.104)
Small firm birth rate	-0.081 (0.165)	**-0.164*** (0.099)	0.021 (0.157)	**-0.171*** (0.099)
Number of small IT service firms	-0.066 (0.066)	**-0.032***** (0.011)	**-0.111***** (0.032)	**-0.033***** (0.008)
Number of small IT service firms per square mile	22.287 (26.187)	4.350 (7.479)	**22.004***** (6.914)	3.771 (3.912)
Constant	0.681 (2.881)	0.118 (2.240)	**-0.666**** (0.278)	**-0.722***** (0.250)
Log likelihood	-583.163	-904.99	-741.301	-1176.113
Overdispersion test (alpha =0)	160.19	398.02	1859.80	2783.6
Vuong test (ZINB vs.NBRM)	3.73	5.02	3.45	4.78
Number of nonzero observations	199	319	214	349
Number of zero observations	102	283	102	283

Note: Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

5.4. Summary

In this chapter, I find that compared to the biopharmaceutical industry, the IT service industry has much larger size, lower barrier for entry, and less degree of spatial concentration. Over the last decade, the industry is most active in IPO and M&A events. However, as in the biopharmaceutical industry, more than half of the deals occurred in a small number of leading regions.

The analysis of the spatial distribution of IT service IPO and M&A events in the catch-up regions provides further evidence of the significant impact of the absolute salary level of scientists, cultural diversity, proximity to venture capital firms, and local entrepreneurial climate on the average NTBF performance in region. The results also suggest that there are noticeable cross-industry differences. There is less evidence of the significance of local scientist job market size in the IT service industry than in the biopharmaceutical industry. The adverse effect of ‘excess’ venture capital and IT service firms is very strong in the IT service industry. Moreover, the industrial specialization is found to be significant and positive only in the IT service industry. Such causal relationships are robust over time, qualitatively similar to the distance-weighted measures and similar to the leading regions.

CHAPTER 6: CONCLUSIONS

This concluding chapter comprises three sections. Section 1 summarizes the main findings of this research. In Section 2, I will discuss the policy implications emerged from the findings. The chapter ends with the discussion of the limitations of this work and future research directions.

6.1. Summary of Main Findings

The empirical analysis presented in this study has endeavored to expand the understanding of the relationship between regional characteristics and the early-stage performance of new technology based firms in the catch-up regions. It introduced a novel measure of NTBF early-stage performance, the event of Initial Public Offering and Merger & Acquisition. It has empirically examined over the past decade how the number of biopharmaceutical or IT service NTBF IPO and M&A events in a catch up region was affected by the local scientist job market conditions, presence of specialized venture capital firms, cultural diversity, industry relevant academic research, industrial structure, and local entrepreneurial climate.

The most striking finding of this study is that the local entrepreneurial climate is a significant determinant of the average NTBF performance in a region. The evidence suggests that the local entrepreneurial climate, measured by the small firm birth rate, is a positive and significant predictor of the number of IPO and M&A events in a region both for the biopharmaceutical and IT service industry. It is also found that its impact has not become weaker over time.

This study also provides some evidence supporting Florida's 'creative class' theory stating that open and diverse regions have distinct advantages in attracting and retaining creative people with unorthodox ideas and therefore are associated with better firm performance. Cultural diversity is found to be significant only in the IT service industry. One explanation is that this study measures a region's cultural diversity by its proportion of foreign born population. Such a measure appears to favor the IT service industry over the biopharmaceutical industry because the former has a relatively higher proportion of foreign born work force.

The empirical findings suggest that among the three variables that measure the local job market conditions for industry-relevant scientists, only the absolute salary level has consistently positive and significant impacts on NTBF performance in both industries. The size of scientist job market has positive and significant impact only in the biopharmaceutical industry. Its impact is insignificant and negative in the IT service industry. The relative salary rate is a negative predictor of the NTBF performance in both industries. Therefore, the results suggest that the "labor market pool" effect is significant only in the highly concentrated industries. In the less concentrated industry like IT service, regions that can provide nationally competitive high-paid jobs for scientists are more likely to attract and retain excellent scientists and technologists and therefore have better NTBF performance.

Proximity to the specialized venture capital firms is found to promote NTBF performance in both industries. However, the results also provide strong evidence of the adverse effects of 'excess' venture capital, particularly in the IT service industry. Similar results regarding the role of VC firms have been previously obtained by Stuart and Sorenson (2003), Acs and Storey (2004), and Venkataraman (2004). The prior literature has pointed out that proximity to many VCs implies that a focal organization may compete against well-financed rivals in local factor market. In addition, the excess amount of venture capital in a region is likely to flow straight to low-quality entrepreneurship.

There is little evidence of the direct effects of industry-related academic research in determining the NTBF performance in catch-up regions. Prior research shows that a strong link exists between academic research and the innovative performance of NTBFs in the biopharmaceutical industry (Audretsch and Stephan 1996; Deeds, Decarolis et al. 1997; McMillan, Narin et al. 2000; Lim 2004). This study finds that in catch-up regions, the academic research variable is not significant either in the biopharmaceutical or IT service industry. The impact of academic research becomes significant only after we take into account the evidence in leading regions, and the impact from both the focal metropolitan area and it neighboring regions.

The study provides weak evidence supporting Porter's (2000) industrial cluster theory stating that industrial specialization and proximity to buyer and suppliers are significant to firm performance. Industrial specialization is found to foster NTBF performance only in the IT service industry. Its impact is negative and insignificant in the biopharmaceutical industry. Proximity to buyer industries is found to have significant but negative impact in the biopharmaceutical industry. Proximity to supplier industry has negative and significant impact in the IT service industry. A possible explanation for the negative impact of industrial specialization and proximity to supplier or buyer industries is that industrial clusters affect the firm formation and post-entry performance differently as suggested by Stuart and Sorenson (2003). In their study of US biotech industry, they found that while new firms are generally formed in the mature industrial clusters, geographic proximity to the competing organization including specialized buyer and suppliers may result in adverse effect on performance.

The noticeable cross-industry difference in the direction, significance level, and size of the effect of the independent variables suggests that industry-specific factors are important in determining the causal relationship between regional factors and NTBF performance. The finding suggest that the IT service industry has a larger size, lower barrier for entry, and lower extent of spatial distribution compared to the biopharmaceutical industry.

6.2. Policy Implications

Over the past decade, the traditional industrial recruitment programs have become less appealing to the state and local governments because of their own exceptionally tight fiscal situations as well as the rising outsourcing practice of many U.S. companies ((NGA 2004). In contrast, fostering "homegrown" NTBFs is becoming a popular policy tool used to achieve regional economic prosperity (Feldman and Francis 2004; Feldman, Francis et al. 2005). While there are a plenty of studies exploring the "best practices" in several high-profile leading high technology clusters, namely Silicon Valley and Boston 128, few efforts have exclusively focused on catch-up regions where a critical mass of industrial cluster has not been established yet. As a result, there is a rising concern that policy initiatives derived from the "best practices" in well-established high-technology regions may be of limited use for catch up regions (Feldman, Francis et al. 2005; Tödtling and Trippl 2005). This dissertation aims to explore the most effective strategies to foster "homegrown" NTBFs in catch up regions by examining the underlying mechanisms through which location affect its average NTBFs performance outside the leading mature clusters. Four distinct veins of policy implications emerge from this research.

First, this study provides strong evidence that promoting local entrepreneurial climate is an effective policy instrument for catch-up regions to foster their NTBF growth. Prior studies have found that entrepreneurship is a strong predictor of the spatial heterogeneity in firm formation. This study found that it is also a key contributor to firm performance. Moreover, the evidence indicates that there is only a modest correlation between local entrepreneurial climate and industrial cluster size. Therefore, this policy tool is particularly relevant for the catch-up regions in which there is a lack of strong industrial foundations. Krueger et al. (2000) pointed out that policy initiatives will increase entrepreneurship only if those initiatives positively influence attitudes and thus influence intentions. They argue that promoting entrepreneurial intentions requires promoting perceptions of both feasibility and desirability of entrepreneurial activities among all strata of society.

Second, this study provides further evidence that the presence of industry specialized venture capital firms is a key determinant of biopharmaceutical and IT service NTBF performance and therefore justifies the public efforts to increase the availability of venture capitals. Financial obstacles to the growth of NTBFs have been the focus of much public policy discussion. Prior study shows that NTBFs typically have poor access to debt because of their highly variable returns, asymmetric information and a lack of collateral (Carpenter and Petersen 2002). Thus, there is a concern that a lack of venture capital may be an important barrier to the development of the high-tech sector. The results of this study suggest that institutional factors that affect the availability and cost of venture capital financing may be an import determinant of the comparative advantage of regions in the growth of small high-tech firms.

Third, this study also provides some evidence that supports the creative-class-based approach proposed by (Florida 2002). This approach posits that regions should attract talents by lowering the barrier to entry for human capital, increasing cultural diversity and regional tolerance, and improving local life amenities.

Finally, the weak impact of academic research on NTBF performance in catch up regions suggests that the quality of academic research determines the extent of its spillover effect. The story of Silicon Valley, Boston 128, and other mature clusters has manifested the role of top universities in fostering the birth and growth of local NTBFs. There are two possible explanations for the insignificance of academic research to NTBF performance in catch-up regions. One is that the academia-industry collaborations tend to be strongest between universities and larger, rather than smaller firms. As pointed out by Storey and Tether (1998), if the prime objective of outreach activities is to generate income, it is almost inevitable that universities give priority to the collaborations with large enterprises because such collaborations are usually more cost-effective and 'prestigious' than those with small and medium firms. However, this study shows that the impact of academic research is significant when the leading regions are included into the model. Thus, this explanation is not strongly supported. The other explanation is that there are few prominent universities or research institutes in catch-up regions so that the private firms don't have strong incentives to collaborate with local academic research efforts. This study provides evidence supporting this argument. The data shows that 7 of the top 20 life science academic departments are located in the leading regions and 13 are in the catch-up regions. Given that catch-up regions account for 95 percent of all the U.S. metropolitan areas, they are associated with a disproportionally low share of high-quality universities. Similarly, in the IT service industry, the catch-up regions only have 11, or 55 percent of top 20 computer science academic departments. In sum, building up local research excellence is the key to enhance the spillover effect of academic research on NTBF performance.

6.3. Limitations and Future Research Directions

The body of work in this research by no means fully answers question about the relationship between location and NTBF performance in the catch-up regions. The limitations of the analysis presented here point to several directions for future studies.

First, the event of IPO and M&A is only a rude measure of the early stage success of NTBF. As mentioned before, this indicator doesn't accounts for the fact that there are substantial variations among the IPO or acquired NTBFs in terms of their individual performance. There is a plenty of evidence that firms that have been issued a successful IPO performed poorly later on. The measure also missed those successful firms whose owners are reluctant to going public or being acquired. Thus, a more appropriate measure of firm performance is needed to obtain more robust results.

Second, this study doesn't decompose the impacts of industry-related entrepreneurial activities and that of general entrepreneurship. This study finds that the overall entrepreneurial climate is a key predictor of average NTBF performance in both the biopharmaceutical and IT service industry. However, it might be the case that the majority of entrepreneurial activities in a region just occurred in these two industries. So, maybe the industry-specific entrepreneurship is the real predictor of NTBF performance. In order to formulate the most effective policies, it is worth to exploring whether industry-specific entrepreneurial activities or just the overall entrepreneurial activities are the true determinant of NTBF performance.

Third, this study doesn't explicitly examine the effect of existing economic development policies on local NTBF performance. A similar analysis at the state level would allow for the addition of policy instruments to the model, which could provide more insights about policy implications.

Finally, the temporal stability analysis in this study is based upon a two-period, ten-year-long time frame. Gartner and Shane ((1995) point out that although the region-level factors may changes dramatically at the societal level over periods greater than 10 years, they change very little over short periods of time. Therefore, the length of time that a longitudinal study is based upon is likely to be a critical factor in itself. The 10-year-long time span may not capture the significant effects of the changes in some regional factors on local NTBF performance.

APPENDIX 1: LIST OF LEADING AND TOP 20 CATCH-UP REGIONS

Table 1.1: List of 'Leading' Regions in the Biopharmaceutical Industry

MSA name	1995 index value	1995 percentile rank	2000 index value	2000 percentile rank
New York-Northern New Jersey-Long Island, NY-NJ-PA	5.998	99	6.133	99
Los Angeles-Long Beach-Santa Ana, CA	3.510	98	2.973	98
Boston-Cambridge-Quincy, MA-NH	1.806	98	1.764	97
San Francisco-Oakland-Fremont, CA	1.673	97	1.707	97
San Jose-Sunnyvale-Santa Clara, CA	1.447	97	0.663	92
Philadelphia-Camden-Wilmington, PA-NJ-DE-MD	1.415	96	1.326	96
San Diego-Carlsbad-San Marcos, CA	1.378	95	1.873	98
Chicago-Naperville-Joliet, IL-IN-WI	1.245	95	1.045	95

Table 1.2: List of Top 20 'Catch-Up' Regions in the Biopharmaceutical Industry

MSA name	1995 index value	1995 percentile rank	2000 index value	2000 percentile rank
St. Louis, MO-IL	1.081	94	0.934	94
Miami-Fort Lauderdale-Miami Beach, FL	0.984	94	1.018	95
Minneapolis-St. Paul-Bloomington, MN-WI	0.820	93	0.963	94
Baltimore-Towson, MD	0.724	92	0.525	88
Detroit-Warren-Livonia, MI	0.722	92	0.524	88
Kansas City, MO-KS	0.688	91	0.550	89
Dallas-Fort Worth-Arlington, TX	0.655	91	0.687	92
Houston-Baytown-Sugar Land, TX	0.590	90	0.605	91
Milwaukee-Waukesha-West Allis, WI	0.529	89	0.278	79
Portland-Vancouver-Beaverton, OR-WA	0.524	89	0.605	91
Atlanta-Sandy Springs-Marietta, GA	0.491	88	0.440	86
Denver-Aurora, CO	0.459	88	0.467	86
Austin-Round Rock, TX	0.427	87	0.331	81
San Antonio, TX	0.426	86	0.467	87
Tampa-St. Petersburg-Clearwater, FL	0.395	86	0.525	89
Washington-Arlington-Alexandria, DC-VA-MD-WV	0.393	85	0.743	93
Portland-South Portland-Biddeford, ME	0.362	85	0.387	84
Raleigh-Cary, NC	0.362	84	0.415	85
Sacramento--Arden-Arcade--Roseville, CA	0.361	84	0.303	81
Seattle-Tacoma-Bellevue, WA	0.361	83	0.550	90

Table 1.3: List of 'Leading' Regions in the IT Service Industry

MSA name	1995 index value	1995 percentile rank	2000 index value	2000 percentile rank
New York-Northern New Jersey-Long Island, NY-NJ-PA	6.138	99	1.663	99
Los Angeles-Long Beach-Santa Ana, CA	2.826	99	0.746	98
Chicago-Naperville-Joliet, IL-IN-WI	2.614	99	0.674	97
Washington-Arlington-Alexandria, DC-VA-MD-WV	2.521	98	0.705	98
San Francisco-Oakland-Fremont, CA	2.512	98	1.153	99
Boston-Cambridge-Quincy, MA-NH	2.182	98	0.749	98
Atlanta-Sandy Springs-Marietta, GA	1.652	97	0.380	95
Dallas-Fort Worth-Arlington, TX	1.569	97	0.300	94
Philadelphia-Camden-Wilmington, PA-NJ-DE-MD	1.567	97	0.469	96
San Jose-Sunnyvale-Santa Clara, CA	1.473	96	0.699	97
Minneapolis-St. Paul-Bloomington, MN-WI	1.340	96	0.385	96
Houston-Baytown-Sugar Land, TX	1.038	96	0.193	91
Seattle-Tacoma-Bellevue, WA	0.936	95	0.239	93
Denver-Aurora, CO	0.892	95	0.196	91
Bridgeport-Stamford-Norwalk, CT	0.876	95	0.579	97

Table 1.4: List of Top 20 'Catch-Up' Regions in the IT Service Industry

MSA name	1995 index value	1995 percentile rank	2000 index value	2000 percentile rank
Detroit-Warren-Livonia, MI	0.862	94	0.254	93
Miami-Fort Lauderdale-Miami Beach, FL	0.837	94	0.313	95
Baltimore-Towson, MD	0.741	94	0.311	95
San Diego-Carlsbad-San Marcos, CA	0.684	94	0.228	92
Trenton-Ewing, NJ	0.665	93	0.878	99
Phoenix-Mesa-Scottsdale, AZ	0.628	93	0.104	85
St. Louis, MO-IL	0.627	93	0.106	86
Tampa-St. Petersburg-Clearwater, FL	0.605	92	0.262	94
Portland-Vancouver-Beaverton, OR-WA	0.600	92	0.124	88
Cleveland-Elyria-Mentor, OH	0.583	92	0.239	93
Boulder, CO	0.479	91	0.401	96
Milwaukee-Waukesha-West Allis, WI	0.478	91	0.229	92
Austin-Round Rock, TX	0.455	91	0.152	89
Pittsburgh, PA	0.436	90	0.107	86
Orlando, FL	0.435	90	0.154	89
Hartford-West Hartford-East Hartford, CT	0.425	90	0.186	90
Manchester-Nashua, NH	0.405	89	0.258	94
Indianapolis, IN	0.400	89	0.118	87
Columbus, OH	0.387	89	0.128	88
Cincinnati-Middletown, OH-KY-IN	0.382	88	0.104	85

APPENDIX 2: INDUSTRIAL SHARE BY STATE, IN 1995, 2000, AND 2004

Table 2.1: Share of Biopharmaceutical Industry by State, 1995, 2000, and 2004

State	Employees in 2004	2004 Share	2000 Share	1995 Share	Dif. 95-04
Alabama	764	0.30	0.22	0.11	0.18
Arizona	1119	0.44	0.48	0.72	-0.28
Arkansas	175	0.07	0.03	0.12	-0.05
California	38636	15.08	13.25	14.78	0.30
Colorado	3052	1.19	1.00	0.86	0.33
Connecticut	9758	3.81	4.50	4.37	-0.56
Delaware	2268	0.89	0.76	0.43	0.45
District of Columbia	10	0.00	0.08	0.00	0.00
Florida	6724	2.63	1.98	1.64	0.98
Georgia	2507	0.98	1.44	0.87	0.11
Hawaii	10	0.00	0.00	0.00	0.00
Idaho	750	0.29	0.08	0.09	0.21
Illinois	19745	7.71	9.28	9.18	-1.47
Indiana	11261	4.40	4.81	4.14	0.25
Iowa	1810	0.71	0.86	1.07	-0.37
Kansas	2485	0.97	0.82	1.00	-0.03
Kentucky	716	0.28	0.16	0.08	0.20
Louisiana	375	0.15	0.22	0.38	-0.23
Maine	1750	0.68	0.45	0.43	0.25
Maryland	4997	1.95	2.02	2.17	-0.22
Massachusetts	8759	3.42	2.94	2.98	0.44
Michigan	7792	3.04	3.87	3.82	-0.78
Minnesota	2396	0.94	0.87	0.92	0.01
Mississippi	1750	0.68	0.76	0.85	-0.17
Missouri	5570	2.17	2.06	2.98	-0.80
Montana	60	0.02	0.03	0.03	-0.01
Nebraska	2079	0.81	0.76	0.78	0.03
Nevada	361	0.14	0.16	0.05	0.10
New Hampshire	790	0.31	0.08	0.09	0.22
New Jersey	35845	13.99	13.35	13.67	0.32
New Mexico	60	0.02	0.13	0.03	-0.01
New York	22380	8.74	8.64	8.86	-0.13
North Carolina	14731	5.75	6.10	5.72	0.03
Ohio	3847	1.50	1.66	1.25	0.25
Oklahoma	750	0.29	0.16	0.13	0.16
Oregon	657	0.26	0.33	0.30	-0.04
Pennsylvania	15033	5.87	5.51	5.52	0.35
Rhode Island	175	0.07	0.08	0.10	-0.03
South Carolina	3327	1.30	1.59	1.09	0.21
South Dakota	60	0.02	0.08	0.13	-0.10
Tennessee	2409	0.94	0.76	1.47	-0.53
Texas	6468	2.53	3.19	3.30	-0.78
Utah	2779	1.08	1.05	0.92	0.16
Vermont	375	0.15	0.00	0.00	0.14

Virginia	3415	1.33	1.45	0.88	0.45
Washington	1615	0.63	0.43	0.49	0.15
West Virginia	1750	0.68	0.76	0.37	0.32
Wisconsin	2004	0.78	0.74	0.81	-0.03
Wyoming	10	0.00	0.00	0.00	0.00

Source:

1. County Business Pattern, Census Bureau, Department of Commerce
2. Some extreme values were verified by Quarterly Census of Employment and Salarys (QCEW) data, Bureau of Labor Statistics (BLS) of the U.S. Department of Labor

Table 2.2: Share of IT Service Industry by State, 1995, 2000, and 2004

State	Employees in 2004	2004 Share	2000 Share	1995 Share	Dif. 95-04
Alabama	17,898	1.25	0.98	0.98	0.27
Alaska	1,194	0.08	0.06	0.03	0.05
Arizona	23,068	1.61	1.51	1.22	0.39
Arkansas	3,560	0.25	0.28	0.43	-0.18
California	244,743	17.07	17.74	15.99	1.08
Colorado	36,863	2.57	3.19	2.67	-0.10
Connecticut	17,549	1.22	1.32	1.51	-0.29
Delaware	4,423	0.31	0.31	0.26	0.05
District of Columbia	10,328	0.72	0.46	0.56	0.16
Florida	59,833	4.17	3.74	4.00	0.17
Georgia	43,315	3.02	3.58	3.52	-0.50
Hawaii	2,496	0.17	0.12	0.13	0.05
Idaho	2,343	0.16	0.15	0.16	0.01
Illinois	52,347	3.65	5.01	4.68	-1.03
Indiana	10,460	0.73	0.90	1.17	-0.44
Iowa	5,310	0.37	0.46	0.90	-0.53
Kansas	9,040	0.63	0.64	0.54	0.09
Kentucky	6,907	0.48	0.47	0.77	-0.29
Louisiana	5,839	0.41	0.35	0.31	0.09
Maine	2,613	0.18	0.14	0.11	0.07
Maryland	59,810	4.17	3.56	3.70	0.47
Massachusetts	69,865	4.87	5.29	5.64	-0.76
Michigan	35,126	2.45	2.14	2.93	-0.48
Minnesota	28,065	1.96	2.19	2.34	-0.38
Mississippi	3,037	0.21	0.15	0.18	0.03
Missouri	24,842	1.73	1.54	1.72	0.01
Montana	2,128	0.15	0.11	0.11	0.04
Nebraska	6,113	0.43	0.45	1.17	-0.75
Nevada	5,931	0.41	0.27	0.14	0.27
New Hampshire	9,229	0.64	0.56	0.47	0.18
New Jersey	63,393	4.42	5.14	4.69	-0.27
New Mexico	4,596	0.32	0.21	0.24	0.08
New York	69,016	4.81	5.24	5.82	-1.00
North Carolina	35,655	2.49	1.94	1.64	0.84
North Dakota	3,768	0.26	0.17	0.10	0.17
Ohio	38,456	2.68	3.00	3.50	-0.82
Oklahoma	9,797	0.68	0.51	0.54	0.14
Oregon	15,687	1.09	1.13	0.95	0.14
Pennsylvania	55,133	3.85	3.87	3.69	0.16
Rhode Island	4,246	0.30	0.19	0.38	-0.08
South Carolina	8,640	0.60	0.49	0.62	-0.02
South Dakota	1,304	0.09	0.09	0.09	0.00
Tennessee	8,819	0.62	0.73	0.84	-0.22
Texas	99,018	6.91	6.64	7.21	-0.30
Utah	13,983	0.98	1.19	1.07	-0.10
Vermont	1,731	0.12	0.17	0.13	-0.01
Virginia	122,421	8.54	7.13	6.91	1.62

Washington	58,386	4.07	3.44	2.00	2.07
West Virginia	2,053	0.14	0.10	0.14	0.00
Wisconsin	13,352	0.93	0.93	1.09	-0.16
Wyoming	529	0.04	0.03	0.02	0.02

Source:

3. County Business Pattern, Census Bureau, Department of Commerce
4. Some extreme values were verified by Quarterly Census of Employment and Salarys (QCEW) data, Bureau of Labor Statistics (BLS) of the U.S. Department of Labor

APPENDIX 3: RESULTS OF REDUCED ZINB MODELS

Table 3.1: Results of Cross-Sectional ZINB Models of Biopharmaceutical IPO and M&A Events between 1996 and 2005

Independent variables	Reduced Model 1	Reduced Model 2	Reduced Model 3
Life scientists job market share	0.403 (0.323)	0.559 (1.927)	0.564 (0.407)
Life scientists absolute salary ratio	1.454 (1.751)		
Life scientists relative salary ratio		-1.006 (16.575)	-0.653 (4.318)
Number of biotech VC firms	**0.330*** (0.181)	0.324 (0.365)	0.312 (0.202)
Square of number of biotech VC firms	-0.049 (0.039)	-0.040 (0.068)	-0.038 (0.058)
Foreign born population share	0.054 (0.043)	0.061 (0.205)	0.061 (0.074)
Life Science Academic research share	0.086 (0.393)	0.094 (0.829)	0.043 (0.712)
Biopharmaceutical industry location quotient	-0.045 (0.073)	-0.014 (0.320)	0.058 (0.052)
Buyer- industry location quotient	1.739 (5.801)	3.021 (44.689)	2.084 (9.097)
Supplier- industry location quotient	-0.053 (0.070)	-0.054 (0.295)	-0.044 (0.104)
Small firm birth rate	7.189 (5.510)	9.590 (36.550)	7.550 (10.158)
Number of small biopharmaceutical firms	**0.162*** (0.091)	**0.147*** (0.081)	**0.180*** (0.092)
Square of number of small biopharmaceutical firms	**-0.005**** (0.002)	**-0.005*** (0.003)	**-0.005**** (0.002)
Number of small biopharmaceutical firms per square mile	57.028 (40.654)	71.128 (134.815)	1.249 (59.221)
Constant	**-3.264*** (1.670)	-1.039 (10.514)	-1.222 (3.032)

Table 3.1 (continued)

Inflation variables	Reduced Model 1	Reduced Model 2	Reduced Model 3
Life scientists job market share	-5.120 (4.186)	-5.613 (6.886)	-5.184* (2.748)
Life scientists absolute salary	-18.615 (11.788)	-25.433 (257.170)	-22.981 (36.033)
Life scientists relative salary	**11.178**** (5.653)	10.128 (32.730)	10.199 (8.826)
Number of biotech VC firms	**-15.958***** (4.232)	-19.548 (43.860)	**-16.543**** (6.659)
Foreign born population share	0.301 (0.212)	0.385 (2.226)	0.362 (0.463)
Life Science academic research share	-2.552 (2.138)	-3.590 (41.816)	-3.292 (9.084)
Biopharmaceutical industry location quotient	-0.075 (0.122)	-0.031 (1.517)	0.015 (0.288)
Buyer- industry location quotient	-3.420 (41.289)	7.388 (494.814)	
Supplier- industry location quotient	-0.336 (0.526)	-0.405 (1.615)	-0.389 (0.349)
Small firm birth rate	-31.623 (32.440)	-33.698 (190.277)	-32.265 (35.665)
Constant	9.264 (10.005)	16.722 (200.923)	14.282 (38.624)
Log likelihood	-157.2407	-158.1929	-152.209
/lnalpha	-3.831	-3.654	-3.499
alpha	0.022	0.026	0.03
Vuong test statistics (zinb vs. nbrm)	2.42	3.06	2.94
Number of observations	160	160	153
Number of nonzero observations	65	65	63
Number of zero observations	95	95	90

Note: Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table 3.2: Results of Two-Period Panel ZINB Models of Biopharmaceutical IPO and M&A Events between 1996 and 2005

Independent variables	**Reduced Model 1**	**Reduced Model 2**	**Reduced Model 3**
Life scientists job market share	0.924*** (0.304)	0.908*** (0.321)	0.893*** (0.291)
Life scientists absolute salary ratio	-0.987 (0.852)		
Life scientists relative salary ratio		-2.095** (0.913)	-2.087** (0.913)
Number of biotech VC firms	0.281 (0.206)	0.272 (0.280)	0.282 (0.262)
Square of number of biotech VC firms	-0.033 (0.031)	-0.034 (0.041)	-0.035 (0.037)
Foreign born population share	0.014 (0.024)	0.010 (0.022)	0.011 (0.023)
Life Science Academic research share	0.002 (0.189)	0.045 (0.166)	0.045 (0.153)
Biopharmaceutical industry location quotient	-0.032 (0.069)	-0.017 (0.075)	0.006 (0.069)
Buyer- industry location quotient	-0.603 (0.701)	-0.630 (0.768)	-0.616 (0.852)
Supplier- industry location quotient	0.041 (0.109)	0.016 (0.112)	0.020 (0.110)
Small firm birth rate	12.399** (5.711)	15.245** (6.338)	14.691** (6.185)
Number of small biopharmaceutical firms	0.091 (0.064)	0.084 (0.082)	0.092 (0.074)
Square of number of small biopharmaceutical firms	-0.003* (0.001)	-0.002 (0.002)	-0.003 (0.002)
Number of small biopharmaceutical firms per square mile	70.289*** (15.307)	58.390*** (14.410)	44.136 (31.180)
Time fixed effect	0.810* (0.442)	1.078** (0.447)	1.047** (0.466)
Constant	-2.199 (1.425)	-0.985 (1.492)	-0.961 (1.568)

Table 3.2 (continued)

Inflation variables	Reduced Model 1	Reduced Model 2	Reduced Model 3
Life scientists job market share	-2.064 (1.390)	-2.200 (1.608)	-2.418 (1.556)
Life scientists absolute salary	-23.381 (14.275)	**-25.021**** (11.833)	**-24.178*** (12.496)
Life scientists relative salary	**7.589*** (4.168)	4.526 (6.345)	2.614 (8.135)
Number of biotech VC firms	-2.485 (1.816)	-3.389 (3.204)	-4.083 (2.884)
Foreign born population share	0.077 (0.116)	0.076 (0.109)	0.065 (0.136)
Life Science academic research share	-7.858 (5.639)	-9.429 (7.921)	-10.930 (11.117)
Biopharmaceutical industry location quotient	0.019 (0.200)	0.092 (0.349)	0.172 (0.338)
Buyer- industry location quotient	-0.528 (2.084)	-0.196 (2.835)	0.134 (4.367)
Supplier- industry location quotient	0.442 (0.352)	0.497 (0.577)	0.588 (0.672)
Small firm birth rate	-1.016 (25.380)	9.358 (40.985)	16.207 (43.622)
Constant	14.078 (10.441)	**17.952**** (8.877)	**18.432*** (9.495)
Log likelihood	-230.38	-226.862	-221.587
/lnalpha	-4.318	-3.831	-3.319
alpha	0.013	0.022	0.036
Vuong test statistics (zinb vs. nbrm)	2.37	2.9	2.83
Number of observations	320	320	310
Number of nonzero observations	89	89	87
Number of zero observations	231	231	223

Note: Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table 3.3: Results of Cross-Sectional ZINB Models of IT Service IPO and M&A Events between 1996 and 2005

Independent variables	Reduced Model 1	Reduced Model 2	Reduced Model 3
Computer scientists job market share	-0.131 (0.163)	-0.060 (0.137)	-0.131 (0.081)
Computer scientists absolute salary ratio	**1.297***** (0.446)		
Computer scientists relative salary ratio		0.333 (0.624)	0.394 (0.600)
Number of IT service VC firms	**0.152***** (0.047)	**0.144***** (0.054)	**0.143***** (0.055)
Square of number of IT service VC firms	**-0.012**** (0.006)	**-0.011**** (0.005)	**-0.011**** (0.005)
Foreign born population share	**0.039*** (0.021)	**0.048***** (0.011)	**0.052***** (0.011)
Computer science academic research share	0.031 (0.028)	0.037 (0.028)	0.045* (0.025)
IT service industry location quotient	**0.476***** (0.089)	**0.502***** (0.094)	**0.500***** (0.086)
Buyer- industry location quotient	-0.006 (0.031)	-0.003 (0.029)	-0.001 (0.025)
Supplier- industry location quotient	**-0.013**** (0.006)	**-0.016***** (0.006)	**-0.016***** (0.006)
Small firm birth rate	**13.034***** (2.063)	**12.445***** (2.416)	**12.449***** (2.436)
Number of small IT service firms	**0.008***** (0.003)	**0.009***** (0.003)	**0.009***** (0.003)
Square of number of small IT service firms	**-0.000**** (0.000)	**-0.000**** (0.000)	**-0.000***** (0.000)
Number of small IT service firms per square mile	0.045 (0.442)	0.309 (0.311)	-0.180 (0.868)
Constant	**-2.594***** (0.633)	**-1.811***** (0.610)	**-1.916***** (0.617)

Table 3.3 (continued)

Inflation variables	Reduced Model 1	Reduced Model 2	Reduced Model 3
Computer scientists job market share	-0.740 (1.108)	-0.803 (1.682)	-1.057 (2.707)
Computer scientists absolute salary	-6.673 (5.090)	**-7.756**** (3.872)	**-9.918*** (5.151)
Computer scientists relative salary	**4.994**** (2.292)	**5.148**** (2.377)	**6.662**** (2.762)
Number of IT service VC firms	0.301 (2.860)	0.431 (2.234)	0.404 (1.580)
Computer science academic research share	1.710 (3.375)	1.772 (2.594)	1.816 (2.160)
IT service industry location quotient	-0.093 (1.002)	-0.108 (0.965)	-0.274 (1.297)
Buyer- industry location quotient	0.018 (0.243)	0.029 (0.175)	0.033 (0.147)
Supplier- industry location quotient	-0.140 (0.104)	-0.136 (0.092)	-0.136 (0.105)
Small firm birth rate	-8.171 (43.467)	-7.564 (34.509)	-6.840 (26.689)
Number of small IT service firms	-0.075 (0.199)	-0.080 (0.151)	-0.081 (0.114)
Number of small IT service firms per square mile	14.698 (43.156)	16.205 (32.210)	27.640 (41.907)
Constant	0.768 (2.355)	1.567 (2.397)	1.098 (2.456)
Log likelihood	-569.712	-572.173	-551.409
/lnalpha	-2.617	-2.564	-2.587
alpha	0.073	0.077	0.075
Vuong test statistics (zinb vs. nbrm)	1.52	1.56	1.62
Number of observations	301	301	290
Number of nonzero observations	199	199	191
Number of zero observations	102	102	99

Note: Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table 3.4: Results of Two-Period Panel ZINB Models of IT Service IPO and M&A Events between 1996 and 2005

Independent variables	Reduced Model 1	Reduced Model 2	Reduced Model 3
Computer scientists job market share	-0.098 (0.103)	-0.020 (0.096)	0.020 (0.106)
Computer scientists absolute salary ratio	**1.899***** (0.554)		
Computer scientists relative salary ratio		**0.880**** (0.370)	**0.857**** (0.366)
Number of IT service VC firms	**0.226***** (0.050)	**0.226***** (0.051)	**0.228***** (0.051)
Square of number of IT service VC firms	**-0.017***** (0.005)	**-0.017***** (0.005)	**-0.017***** (0.005)
Foreign born population share	0.014 (0.010)	**0.029***** (0.009)	**0.029***** (0.009)
Computer science academic research share	0.010 (0.030)	0.011 (0.028)	0.007 (0.028)
IT service industry location quotient	**0.451***** (0.102)	**0.519***** (0.112)	**0.527***** (0.117)
Buyer- industry location quotient	-0.020 (0.013)	-0.018 (0.013)	-0.019 (0.013)
Supplier- industry location quotient	-0.011 (0.010)	-0.016 (0.010)	**-0.017*** (0.010)
Small firm birth rate	**9.952***** (1.607)	**7.669***** (1.617)	**7.488***** (1.609)
Number of small IT service firms	**0.004***** (0.001)	**0.004***** (0.001)	**0.004***** (0.001)
Square of number of small IT service firms	**-0.000***** (0.000)	**-0.000***** (0.000)	**-0.000***** (0.000)
Number of small IT service firms per square mile	-0.082 (0.193)	0.200 (0.217)	0.058 (0.201)
Time fixed effect	-0.187 (0.136)	**-0.405***** (0.134)	**-0.401***** (0.142)
Constant	**-2.815***** (0.809)	**-2.070***** (0.540)	**-2.000***** (0.543)

Table 3.4 (continued)

Inflation variables	**Reduced Model 1**	**Reduced Model 2**	**Reduced Model 3**
Computer scientists job market share	-0.060	-0.014	0.033
	(0.262)	(0.259)	(0.254)
Computer scientists absolute salary	-2.985	**-4.353****	**-5.368****
	(2.187)	(2.198)	(2.372)
Computer scientists relative salary	**3.726****	**4.062*****	**4.691*****
	(1.485)	(1.342)	(1.475)
Number of IT service VC firms	0.244	0.286	0.324
	(0.531)	(0.498)	(0.510)
Foreign born population share	0.015	0.023	0.019
	(0.032)	(0.033)	(0.035)
Computer science academic research share	-0.063	-0.068	-0.061
	(0.233)	(0.224)	(0.224)
IT service industry location quotient	**1.506***	**1.481***	**1.585****
	(0.803)	(0.762)	(0.808)
Buyer- industry location quotient	**-0.272***	**-0.285***	**-0.295****
	(0.141)	(0.146)	(0.144)
Supplier- industry location quotient	**-0.176***	**-0.177***	**-0.195***
	(0.106)	(0.105)	(0.113)
Small firm birth rate	**-17.697***	**-19.280****	**-19.332***
	(9.614)	(9.555)	(9.897)
Number of small IT service firms	**-0.049*****	**-0.049*****	**-0.051*****
	(0.016)	(0.014)	(0.014)
Number of small IT service firms per square mile	3.900	3.886	5.037
	(7.716)	(7.264)	(7.476)
Constant	0.200	1.166	1.241
	(2.238)	(1.863)	(1.910)
Log likelihood	-864.266	-867.916	-851.304
/lnalpha	-2.479	-2.555	-2.572
alpha	0.084	0.078	0.076
Vuong test statistics (zinb vs. nbrm)	2.60	2.1	2.11
Number of observations	602	602	591
Number of nonzero observations	319	319	314
Number of zero observations	283	283	277

Note: Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

REFERENCES

Acs, Z. J. and D. J. Storey (2004). "Introduction: Entrepreneurship and Economic Development." Regional Studies **38**(8): 871-877.

Audretsch, D. and M. Feldman (1996). "R&D spillovers and the geography of innovation and production." American Economic Review **63**(3): 630-640.

Audretsch, D. B. and E. E. Lehmann (2005). "Does the Knowledge Spillover Theory of Entrepreneurship hold for regions?" Research Policy **34**(8): 1191-1202.

Audretsch, D. B. and P. Stephan (1996). "Company-scientist locational linkages: the case of biotechnology." American Economic Review **86**(3): 641-652.

Bee, E. (2003). "Knowledge Networks and Technical Invention in America's Metropolitan Areas: A Paradigm for High-Technology Economic Development." Economic Development Quarterly **17**(2): 115-131.

Benninga, S., M. Helmantel, et al. (2005). "The timing of initial public offerings." Journal of Financial Economics **75**(1): 115-132.

Bollinger, L., K. Hope, et al. (1983). "A review of literature and hypotheses on new technology-based firms." Research Policy **12**(1): 1-14.

Brau, J. C., R. A. Brown, et al. (2004). "Do venture capitalists add value to small manufacturing firms? An empirical analysis of venture and nonventure capital-backed." Journal of Small Business Management **42**(1): 78-92.

Carpenter, R. E. and B. C. Petersen (2002). "Capital market imperfections, high-tech investment, and new equity financing." Economic Journal **112**(477): F54-F72.

Colombo, M. G. and L. Grilli (2005). "Founders' human capital and the growth of new technology-based firms: A competence-based view." Research Policy **34**(6): 795-816.

Deeds, D., D. Decarolis, et al. (1997). "The impact of firm-specific capabilities on the amount of capital raised in an initial public offering: Evidence from the biotechnology industry." JOURNAL OF BUSINESS VENTURING **12**(1): 31-46.

DeVol, R. C. (1999). America's High-Tech Economy: Growth, Development, and Risks for Metropolitan Areas, Milken Institute.

Feldman, M. and D. B. Audretsch (1999). "Innovation in cities: Science-based diversity, specialization and localized competition." European Economic Review **43**(2): 409-429.

Feldman, M. and J. L. Francis (2004). "Homegrown solutions: Fostering cluster formation." Economic Development Quarterly **18**(2): 127-137.

Feldman, M., J. L. Francis, et al. (2005). "Creating a cluster while building a firm: Entrepreneurs and the formation of industrial clusters." Regional Studies **39**(1): 129-141.

Florida, R. (2002a). "The Economic Geography of Talent." Annals of the Association of American Geographers **92**(4): 743-755.

Florida, R. (2002b). The Rise of The Creative Class: And How It's Transforming Work, Leisure, Community, & Everyday life. NY: New Yorks, Basic Books.

Florida, R. (2003). Entrepreneurship, Creativity, and Regional Economic Growth. The Emergence of Entrepreneurship Policy: Governance, Start-Ups, and Growth in the U.S.Knowledge Economy. D. M. Hart. Cambridge, Cambridge University Press: 39-58.

Folta, T. B., A. C. Cooper, et al. (2006). "Geographic Cluster Size and Firm Performance." Journal of Business Venturing **21**: 217-242.

Gartner, W. B. and S. A. Shane (1995). "Measuring entrepreneurship over time." Journal of Business Venturing **10**(4): 283-301.

Glaeser, E. L., G. Dumais, et al. (2002). "Geographic Concentration as a Dynamic Process." Review of Economics and Statistics **84**(2): 193-204.

Glaeser, E. L., H. Kallal, et al. (1992). "Growth in Cities." Journal of Political Economy **100**(6): 1126-1152.

Guo, R., B. I. Lev, et al. (2005). "The Valuation of Biotech IPOs." Journal of Accounting, Auditing & Finance **20**(4): 188-244.

Hecker, D. E. (2005). "High technology employment: a NAICS-based update." Monthly Labor Review(July): 57-72.

Helwege, J. and N. Liang (2004). "Initial public offerings in hot and cold markets." Journal of Financial and Quantitative Analysis **39**(3): 541-569.

Jaffe, A., M. Trajtenberg, et al. (1993). "Geographic Localization of Knowledge Spillovers as Evidenced by Patent citations." The quarterly Journal of Economics 577-98.

Ketelhohn, N. W. (2002). The Role of Clusters as Sources of Dynamic Externalities. Graduate School of Business, Harvard University.

Krueger, N. F., M. D. Reilly, et al. (2000). "Competing Models of Entrepreneurial Intentions." Journal of Business Venturing **15**: 411-432.

Krugman, P. (1991). "Increasing returns and economic geography." Journal of Political Economy **99**(3): 483-499

Lee, S. Y., R. Florida, et al. (2004). "Creativity and Entrepreneurship: A Regional Analysis of New Firm Formation." Regional Studies **38**(8): 879 - 891

Lim, K. (2004). "The relationship between research and innovation in the semiconductor and pharmaceutical industries (1981-1997) " Research Policy **33**(2): 287-321.

Malmberg, A. and D. Power (2005). On the role of global demand in local innovation processes Rethinking Regional Innovation and Change: Path Dependency or Regional Breakthrough? G. Fuchs and P. Shapira. New York, NY, Springer.

Marshall, A. (1920). Principles of Economics. London, MacMillan.

McMillan, S. G., F. Narin, et al. (2000). "An analysis of the critical role of public science in innovation: the case of biotechnology " Research Policy, **29**(1): 1-8

Motohashi, K. (2005). "University-industry collaborations in Japan: The role of new technology-based firms in transforming the National Innovation System." Research Policy **34**(5): 583-594.

NGA (2004). A Governor's Guide to Building State Science and Technology Capacity, The National Governors Association.

NGA (2004). A Governor's Guide to Strengthening State Entrepreneurship Policy, The National Governors Association.

Ottaviano, G. I. P. and G. Peri (2004). "The Economic Value of Cultural Diversity: Evidence from US Cities." NBER Working Paper No. 10904.

Porter, M. E. (1990). The Competitive Advantage of Nations. New York, The Free Press.

Porter, M. E. (1998). "Clusters And The New Economics Of Competition." Harvard Business Review **76**(6).

Porter, M. E. (2000). "Location, Competition and Economic Development: Local Clusters in a Global Economy." Economic Development Quarterly **14**(1): 15-34.

Porter, M. E. (2003). "the Economic Performance of Regions." Regional Studies **37**(6&7): 549-578.

Ritter, J. R. and I. Welch (2002). "A Review of IPO Activity, Pricing, and Allocations." Journal of Finance **57**(4): 1795-1828.

Saxenian, A. (1999). Silicon Valley's New Immigrant Entrepreneurs. San Francisco, Public Policy Institute of California.

Smith, S. W. (2004). Innovation and Globalization in Four High-Technology Industries in the United States: One Size Does Not Fit All, Harvard University.

Sorenson, O. and P. G. Audia (2000). "The social structure of entrepreneurial activity: Geographic concentration of footwear production in the United States, 1940-1989 " American Journal of Sociology **106**(2): 424-461.

Sorenson, O. and T. E. Stuart (2001). "Syndication networks and the spatial distribution of venture capital investments " AMERICAN JOURNAL OF SOCIOLOGY **106**(6): 1546-1588.

Storey, D. J. and B. S. Tether (1998). "New technology-based firms in the European union: an introduction." Research Policy **26**(9): 933-946.

Storey, D. J. and B. S. Tether (1998). "Public policy measures to support new technology-based firms in the European Union." Research Policy **26**(9): 1037-1057.

Stuart, T. and O. Sorenson (2003). "The geography of opportunity: spatial heterogeneity in founding rates and the performance of biotechnology firms." Research Policy **32**(2): 229-253.

Stuart, T. E. and O. Sorenson (2005). Social Networks and Entrepreneurship. Handbook of Entrepreneurship Research. R. Agarwal, S. Alverez and O. Sorenson. New York, Springer**:** 233-252.

Thornton, P. H. and K. H. Flynn (2003). Entrepreneurship, Networks, and Geographies. Handbook of Entrepreneurship Research. Z. J. Acs and D. B. Audretsch. Great Britain, 2003 Kluwer Law International**:** 401-433.

Tödtling, F. and M. Trippl (2005). "One size fits all?: Towards a differentiated regional innovation policy approach." Research Policy **34**(8): 1203-1219.

Venkataraman, S. (2004). "Regional transformation through technological entrepreneurship " Journal of Business Venturing **19**: 153-167.

Walcott, S. (2000). Defining and Measuring High Technology in Georgia. Report of the Fiscal Research Program.

Zhang, J. and N. Patel (2005). The Dynamics of California's Biotechnology Industry. San Francisco, CA, Public Policy Institute of California.